Communication Equilibria with Partially Verifiable Types*

Françoise Forges † Frédéric Koessler ‡

THEMA (CNRS, UMR 7536)
Université de Cergy-Pontoise
33, Boulevard du Port
95011 Cergy-Pontoise, France

March 6, 2003

Abstract

This paper studies the set of equilibria that can be achieved by adding general communication systems to Bayesian games in which some information can be certified or, equivalently, in which players' types are partially verifiable. Given the information that players are able to certify, we characterize outcome-equivalent canonical equilibria for which generalized versions of the revelation principle are valid. Communication equilibria and associated canonical representations are obtained as special cases when no information can be certified.

Keywords: Bayesian game; Communication equilibrium; Information certification; Revelation principle; Verifiable types.

JEL Classification: C72; D82.

*We thank Jean-François Mertens, Regis Renault, and seminar participants at Nanterre and Paris (Game Theory Seminar at Institut Henri Poincaré) for helpful comments and discussions.

† *E-mail:* `Francoise.Forges@eco.u-cergy.fr`.

‡ *E-mail:* `Frederic.Koessler@eco.u-cergy.fr`.

1 Introduction

Since the pioneering work of Aumann (1974) on correlated equilibria and Crawford and Sobel's (1982) analysis of cheap talk games, the introduction of communication possibilities into the analysis of interactive decision situations has been commonplace in a whole host of applied and theoretical researches (for some recent references, see, e.g., Aumann and Hart, 2003, Baliga and Morris, 2002, Battaglini, 2002, Ben-Porath, 2003, Gerardi, 2003, Krishna and Morgan, 2002, Urbano and Vila, 2002, and Wolinsky, 2002). Such analyses are motivated by the fact that when individuals can talk with each others before choosing their final payoff-relevant actions, they may be able to share information and/or agree on compromises, and then reach outcomes that differ from those of the standard Nash equilibrium solution concept. For example, a correlated equilibrium of a strategic form game is a Nash equilibrium of some extension of the game where players receive private, "extraneous" and possibly correlated signals before the beginning of the original game. Such a solution concept is appropriate to characterize the set of all equilibrium outcomes achievable in one-shot complete information games with cheap and non-binding communication.

With the exception of some specific applications discussed below, the literature on communication games and the various extensions of the correlated equilibrium to incomplete information typically relied on the assumption that the set of reports available to a player does not depend on his private information.[1] On the contrary, our starting point in this paper is to consider that information which is transmitted might be certifiable or provable by its sender, or verifiable by its receiver. Said differently, we assume that the set of all possible arguments an individual is able to present may vary with his actual state of knowledge. For example, reports may consist of written documents or direct physical observations which are not possible to forge unless their information contents are true.[2] Alternatively, in economic or legal interactions there may be penalties for perjury, false advertising and warranty violations, or accounting principles that impose limits on what is possible to disclose. Requiring traders in a exchange economy to deposit collateral for each order (as, e.g., in Forges, Mertens, and Vohra, 2002) also implies that their types are partially verifiable because traders are not able to over-report their initial endowments.[3] Finally, an individual's ability to manipulate and misrepresent information may be limited due to psychological reasons (e.g., observable emotions such as blushing, stress and discomfort, or a strong taste for honesty that cannot adequately be represented by standard preferences, as in Alger and Ma, 2003, and Alger and Renault, 2002). The purpose of this paper is precisely to study in a general and tractable framework the

[1]For an overview, see, e.g., Farrell and Rabin (1996), Forges (1993), and Myerson (1994).

[2]For instance, disclosures of knowledge generated by R&D may be knowledge-dependent in the sense that an informed firm cannot disclose more knowledge than it has (see, e.g., d'Aspremont, Bhattacharya, and Gérard-Varet, 2000).

[3]Similarly, the type of a budget-constrained buyer may be partially verifiable if the seller can ask him to post a bond equal to his reported budget (as, e.g., in Che and Gale, 2000).

effects of adding communication systems to incomplete information games in which players' types are certifiable, provable, or partially verifiable, and to characterize canonical representations and equilibria of such extended communication games.

Our approach combines three areas of research. The first relates to the notion of communication equilibrium (Forges, 1986; Myerson, 1982, 1986). A communication equilibrium is an extension of Aumann's (1974) correlated equilibrium to the framework of incomplete information, which allows players not only to receive recommendations from a mediator, but also to send signals to share some privately held information. We extend this solution concept by allowing players to send certified information into the communication system, in addition to cheap talk signals. This means that the set of reports available to some type of a player can differ from the set of reports available to another type of the *same* player. The associated set of equilibrium outcomes, called *certification equilibrium* outcomes, is larger than the set of communication equilibrium outcomes but it depends on the certifiability configuration (i.e., on the set of exogenously certifiable types in the game) and on the number of communication stages. Indeed, in general, the problem of adverse selection crucially depends on the information that players are able to certify and on the number of arguments they can present.

The second area of research related to our work is the economic literature dealing with strategic information revelation, initiated by Grossman (1981), Grossman and Hart (1980) and Milgrom (1981), which investigates the amount of information voluntarily transmitted when individuals are required to make only truthful—but possibly very vague—disclosures.[4] This literature has particular insight in oligopoly theory (see, e.g., Okuno-Fujiwara, Postlewaite, and Suzumura, 1990), finance (see, e.g., Shin, 2003), and law (see, e.g., Shin, 1994, 1998). The accounting literature has also placed considerable emphasis on games with strategic information revelation (see Dye, 2001, Verrecchia, 2001 and references therein). Contrary to those previous contributions we consider a general game-theoretical framework allowing private, stochastic, repeated, and mediated information revelation, and we do not require players' types to be independent.

Finally, our work is related to the literature on mechanism design with partially verifiable information (Bull and Watson, 2002, Deneckere and Severinov, 2001, Green and Laffont, 1986). This literature, which is restricted to the implementation of an exogenous social choice function, studies the validity of the standard revelation principle when the set of available reports of perfectly informed agents varies with the true state of the world. In particular, several conditions have been identified in which the revelation principle fails because the set of reports agents are able to send to the principal is type-dependent. The difference with those previous contributions is that our revelation principle is defined for (n-person) games of incomplete information and allows, given any fixed set of certifiable information, a fully equivalent

[4]For more recent references, see, e.g., Glazer and Rubinstein (2001), Koessler (2002, 2003), Lipman and Seppi (1995), Seidmann and Winter (1997), and Wolinsky (2003).

equilibrium characterization. This is performed by providing canonical representations of communication systems and reporting correspondences which are compatible with the original certification possibilities. In particular, since the set of canonical equilibrium outcomes contains all non-canonical equilibrium outcomes but does not allow to achieve more, our representations can be used without loss of generality to maximize, e.g., a particular player's payoff or a welfare function.

At first sight, it is not clear how to define a canonical communication system and equilibrium when the set of available reports may vary with the state of knowledge of each player. Yet, once the certifiability configuration generated by the original profile of reporting correspondences is well characterized, the canonical representation we propose is simple: players are only required to present, in a one-stage game, the most informative certificate concerning their type to a mediator and to make a cheap talk claim about their type. Then, once the mediator has received a report in this "canonical" space from each player he makes private recommendations to the players. We show that there is no loss of generality in focusing on such representations and on equilibria where players reveal their true type and follow the recommendations of the mediator.[5] If the original set of possible communication systems is restrained to one-period communication systems where players can only present one verifiable argument, we also provide a sufficient condition on the reporting correspondences which maintains the outcome equivalence between the associated certification equilibria and canonical certification equilibria. Finally, we construct an equilibrium characterization in which the set of available reports of each player simply consists in a set of possible types, where it is implicitly assumed that a report of a type is associated with all available certificates compatible with this type. This construction can be done without loss of generality when the set of reports available in the communication system can be restrained by the mediator and in games with transferable utility. Various illustrating examples are examined along the exposition.

The paper is organized as follows. In Section 2 we present our general framework and some preliminary definitions. Canonical representations and generalized versions of the revelation principle for Bayesian games are analyzed in Section 3. We conclude in Section 4. The Appendix contains the proofs.

[5]In particular, the set of certification equilibrium outcomes has, as the set of correlated and communication equilibrium outcomes, a simple and tractable mathematical structure: it is a convex polyhedron.

2 General Framework and Definitions

2.1 Bayesian Games and Communication Systems

We represent an interactive decision situation under asymmetric information by a (finite) *Bayesian game*

$$G = \langle N, (A_i)_{i \in N}, (T_i)_{i \in N}, p, (u_i)_{i \in N} \rangle,$$

where $N = \{1, \ldots, n\}$ is the set of players, A_i is player i's set of possible actions, T_i is player i's set of possible types, $p \in \Delta(T)$ is a common prior probability distribution over the set of type profiles $T = \prod_{i \in N} T_i$, and $u_i : A \times T \to \mathbb{R}$ is player i's state dependent payoff (utility) function, where $A = \prod_{i \in N} A_i$ is the set of action profiles. Let $p(t_i) \equiv \sum_{t_{-i} \in T_{-i}} p(t_i, t_{-i})$ be the prior probability that player i's type is t_i.[6] We assume without loss of generality that $p(t_i) > 0$ for all $i \in N$ and $t_i \in T_i$. Let $p(t_{-i} \mid t_i) \equiv \frac{p(t)}{p(t_i)}$ be the subjective probability that player i assigns to the event that t_{-i} is the actual profile of the other players' types if his own type is t_i.[7]

To allow players to communicate before choosing an action in the Bayesian game G, we introduce a communication system (or mediator) that helps players to share information and to coordinate their actions.[8] As usual, in a game with communication players exchange messages conditionally on past messages and on their own type before choosing their actions. However, contrary to previous work related to cheap talk communication and to the various extensions of the correlated equilibrium to incomplete information, we assume that players are able to certify some of their information. In other words, the set of available reports may be type-dependent, which implies that a report may have some pure informational content which does not depend on any particular equilibrium.

Formally, a (finite) *communication system* given the set of players, N, and the set of possible type profiles, T, is denoted by

$$c = \langle (R_i)_{i \in N}, (S_i)_{i \in N}, (M_i)_{i \in N}, K, (\nu^k)_{k=0,1,\ldots,K} \rangle.$$

The positive integer K is the number of *communication periods*. For each player i, $R_i : T_i \to \mathcal{R}_i$ is a *reporting correspondence* that determines the set $R_i(t_i)$ of *type-dependent inputs* available to player i of type $t_i \in T_i$, i.e., the set of *reports* that player i can send out into the communication system in each period if his actual type is t_i, and $\mathcal{R}_i \equiv \bigcup_{t_i \in T_i} R_i(t_i)$ is the set of all reports the communication system can receive from player i in each period. The set S_i is the set of *type-independent inputs* available to player i, i.e., the set of *cheap talk signals* that player i can send

[6] For any variable, we denote its profile over all agents except that of player i by the corresponding letter with subscript $-i$.

[7] We do *not* assume that every type profile has non-zero probability.

[8] Players have no ability to sign any contract or binding agreement. Hence, our approach is strictly non-cooperative.

out into the communication system in each period. The set M_i is the set of *outputs* for player i, i.e., the set of all *messages* that player i can privately receive from the communication system in each period. Let $\mathcal{R} = \prod_{i \in N} \mathcal{R}_i$, $S = \prod_{i \in N} S_i$, and $M = \prod_{i \in N} M_i$. (Observe that, a priori, the elements of $\mathcal{R}$, S and M have no semantic content.) In period 0, each player i privately receives from the communication system an *initial output* $m_i^0 \in M_i$ according to the probability distribution $\nu^0 \in \Delta(M)$. Then, at the end of each communication period $k \in \{1, \ldots, K\}$, after all inputs up to that period have been received by the communication system, the *transition probability*

$$\nu^k : M^k \times \mathcal{R}^k \times S^k \to \Delta(M),$$

chooses the outputs as a function of past outputs and past and present inputs. That is, $\nu(m^k \mid m^0, m^1, \ldots, m^{k-1}, r^1, \ldots, r^k, s^1, \ldots, s^k)$ is the conditional probability that $m^k = (m_1^k, \ldots, m_n^k) \in M$ are the messages privately received by the various players at the end of period k given the sequence of vectors of past outputs $(m^0, m^1, \ldots, m^{k-1}) \in M^k$, past and present type-dependent inputs $(r^1, \ldots, r^k) \in \mathcal{R}^k$, and past and present type-independent inputs $(s^1, \ldots, s^k) \in S^k$.

2.2 Extended Bayesian Games and Certification Equilibria

Given a communication system c, one can define the extension G_c of G as the new game obtained by adding c to G. Such a communication game proceeds as follows. In period 0 every player i receives a confidential initial output $m_i^0 \in M_i$, where $m^0 = (m_i^0)_{i \in N}$ is distributed according to ν^0, and is privately informed about his type $t_i \in T_i$, where $t = (t_i)_{i \in N}$ is distributed according to p. Then, at the beginning of each period $k \in \{1, \ldots, K\}$ he sends a confidential input $(r_i^k, s_i^k) \in R_i(t_i) \times S_i$ to the communication system. At the end of each period $k \in \{1, \ldots, K\}$, he receives a confidential output $m_i^k \in M_i$ from the communication system, where $m^k = (m_i^k)_{i \in N}$ is conditionally distributed according the ν^k. Finally, after the last communication period (in period $K+1$, which corresponds to the action phase) he chooses an action $a_i \in A_i$ and is rewarded according to his utility function u_i.

A behavioral strategy for player i in G_c is a tuple $((\sigma_i^k)_{k=1,\ldots,K}, \delta_i)$ where for all $k \in \{1, \ldots, K\}$,

$$\sigma_i^k : M_i^k \times \mathcal{R}_i^{k-1} \times S_i^{k-1} \times T_i \to \Delta(\mathcal{R}_i \times S_i),$$

is player i's communication strategy in period k satisfying $\sigma_i^k(r_i^k, s_i^k \mid \cdot, t_i) = 0$ whenever $r_i^k \notin R_i(t_i)$, and

$$\delta_i : M_i^{K+1} \times \mathcal{R}_i^K \times S_i^K \times T_i \to \Delta(A_i),$$

is player i's strategy in the action phase. A profile of behavioral strategies is denoted by $(\sigma, \delta) = (\sigma_i, \delta_i)_{i \in N}$, where $\sigma_i = (\sigma_i^k)_{k=1,\ldots,K}$. Such a strategy profile in G_c generates an *outcome* $\mu : T \to \Delta(A)$ (i.e., a conditional probability distribution over A for each type profile $t \in T$) and an expected payoff $\sum_{t\in T} p(t) \sum_{a\in A} \mu(a \mid t)\, u_i(a,t)$ for each player i.[9] As usual, a (Bayesian) Nash equilibrium of the communication game G_c is a strategy profile (σ,δ) such that no player can strictly increase his expected payoff by unilaterally deviating from his strategy. The outcome generated by a Nash equilibrium of G_c is called an *equilibrium outcome* of G_c.[10]

Definition 1 A *certification equilibrium* of G is a Nash equilibrium of the extended game G_c obtained by adding a communication system c to G.

It can be shown[11] that the set of all certification equilibrium outcomes, denoted by $\mathcal{E} \subseteq [\Delta(A)]^T$, obtained when considering *all* possible communication systems (in particular, all possible reporting correspondences), coincides with the set of Nash equilibrium outcomes of the extended games obtained by adding a one-period communication system ($K = 1$) without initial output (ν^0 is degenerated), without type-independent input (S is a singleton), satisfying $M = A$, $R_i(t_i) = \{t_i\}$ for all $i \in N$ and $t_i \in T_i$, and in which every player follows the recommendation of the mediator. That is, a certification equilibrium outcome is simply characterized by a recommendation $\mu : T \to \Delta(A)$ satisfying

$$\sum_{t_{-i}\in T_{-i}} p(t_{-i} \mid t_i) \sum_{a\in A} \mu(a \mid t)\, u_i(a;t) \geq \sum_{t_{-i}\in T_{-i}} p(t_{-i} \mid t_i) \sum_{a\in A} \mu(a \mid t)\, u_i(a_{-i}, d_i(a_i);t),$$

for all $i \in N$, $t_i \in T_i$, and $d_i : A_i \to A_i$. The intuition of this equivalent characterization is very simple. Starting with any certification equilibrium, the mediator first simulates the sequence of signals and reports (inputs) that would have been sent by the players and the sequence of messages (outputs) that would have been received by the players given the type profile under the original equilibrium. Then, he computes the actions that would have been chosen by the players as a function of the type profile and the sequence of inputs and outputs. Finally, he privately recommends each player to choose the associated action. Clearly, if a player has an incentive to deviate from the recommendation of the mediator, then the strategy profile of the original communication game was not an equilibrium.

The previous observation can be interpreted as a form of "revelation principle": any certification equilibrium is outcome equivalent to a "truthful certification equilibrium". However, the set of "truthful certification equilibria" generated in this way is much too large for the result to be interesting, and is not appropriate for most applications. Indeed, players may have the right to remain silent or to present only vague arguments, whereas in some certification equilibria they are compelled to

[9] That is, if for all $(m,r,s) \in M^K \times \mathcal{R}^K \times S^K$ we denote by $h(m,r,s \mid m^0,t)$ the probability distribution over $M^K \times \mathcal{R}^K \times S^K$ generated by (σ,δ) in G_c given $m^0 \in M$ and $t \in T$, then $\mu(a \mid t) = \sum_{m^0\in M} \nu^0(m^0) \sum_{(m,r,s)\in M^K\times\mathcal{R}^K\times S^K} h(m,r,s \mid m^0,t)\delta(a \mid m^0,m,r,s,t)$.

[10] We consider equilibrium outcomes rather than equilibrium strategies because the dimension of strategy sets depends on the underlying communication system. By contrast, equilibrium outcomes are always in $[\Delta(A)]^T$.

[11] The formal proof is a simplified version of the Proof of Theorem 1.

reveal their type to the mediator even if they have no incentive to do so. A simple illustration is provided in Example 1. On the other hand, in some environments players may have only limited ability to certify claims. Accordingly, when certifiability possibilities are given and only partial, it is not appropriate to consider a communication system with $R_i(t_i) = \{t_i\}$ for all $i \in N$ and $t_i \in T_i$ because what is certified with such a communication system might not be certifiable with the original set of available reports.

For those reasons we define certification equilibria that can be obtained only with a specified profile of available type-dependent inputs, i.e., with communication systems where the reporting correspondences $R = (R_i)_{i \in N}$ are given. Such communication systems are called *R–communication systems*. Of course, if the set of available inputs does not depend on players' types then the set of associated equilibria is, by definition, the set of communication equilibria.

Definition 2 A *R–certification equilibrium* of G is a Nash equilibrium of the extended game G_c obtained by adding a R–communication system c to G. A *communication equilibrium* is a R–certification equilibrium where $R_i(t_i) = R_i(t_i')$ for all t_i, $t_i' \in T_i$ and $i \in N$.

We denote by $\mathcal{E}(R)$ the set of R–certification equilibrium outcomes and by $\mathcal{E}_0$ the set of communication equilibrium outcomes. Clearly, we have $\mathcal{E}_0 \subseteq \mathcal{E}(R) \subseteq \mathcal{E}$ for every profile of reporting correspondences R, and all these sets are convex (thanks to the preliminary lottery ν^0). As shown in the following example these inclusions may be strict.

Example 1 Consider a consumer whose endowments depend on two equally likely types, t^1 and t^2, which are private information to the consumer. There are two commodities. In state t^1 (t^2, resp.) the consumer's endowment is $(10, 0)$ ($(0, 10)$, resp.). A government can choose to deduct taxes of twenty per cent either on commodity 1 (action a^1) or on commodity 2 (action a^2). If each unit of commodity provides a utility of one to the consumer and to the government, this situation can be represented by the Bayesian game of Figure 1 on the next page. In this game the set of communication equilibrium outcomes and the set of R–certification equilibrium outcomes coincide whenever player 1 can remain silent, i.e., whenever $\bigcap_{t \in T} R_1(t) \neq \emptyset$: they are characterized by $\mu(a^2 \mid t^2) = 1 - \mu(a^1 \mid t^1)$. Hence, the only associated vector of expected payoffs is $(9, 1)$. The set of *all* certification equilibrium outcomes is however *strictly larger* since it is the set of outcomes satisfying $\mu(a^2 \mid t^2) \geq 1 - \mu(a^1 \mid t^1)$. In particular, the perfectly revealing recommendation induces such an equilibrium outcome with the vector of expected payoffs $(8, 2)$.

In the following section we introduce canonical communication systems and equilibria given some specified profile of reporting correspondences $R = (R_i)_{i \in N}$ in order to obtain a simple and equivalent characterization of the set of all R–certification equilibrium outcomes.

	a^1	a^2
t^1	$(8,2)$	$(10,0)$
t^2	$(10,0)$	$(8,2)$

Figure 1: Bayesian Game of Example 1.

3 Canonical Representations

3.1 Certifiability Configuration and Canonical Communication Systems

As noted earlier, the inputs in a communication system have no semantic content. In order to capture certification possibilities in a canonical way, we represent certifiable information as events of the state space. More precisely, we define a *certifiability configuration* as an n-tuple of collections of sets of types, $\mathcal{Y} = (\mathcal{Y}_i)_{i\in N}$, where an element $y_i \in \mathcal{Y}_i \subseteq 2^{T_i}\backslash\{\emptyset\}$ is a *certificate* (certifiable event) concerning player i's type. For all $i \in N$ and $t_i \in T_i$ we assume that there exists $y_i \in \mathcal{Y}_i$ such that $t_i \in y_i$.[12] The set of events that player i of type t_i is able certify concerning his type is the set of certificates containing t_i and is denoted by $Y_i(t_i) \equiv \{y_i \in \mathcal{Y}_i : t_i \in y_i\}$. Hence, a certifiability configuration $\mathcal{Y} = (\mathcal{Y}_i)_{i\in N}$ can equivalently be viewed as a profile of reporting correspondences $R = Y = (Y_i)_{i\in N}$. The *closure* of a certifiability configuration Y is the certifiability configuration $\overline{Y} = (\overline{Y}_i)_{i\in N}$ where for all $i \in N$ and $t_i \in T_i$, $\overline{Y}_i(t_i)$ is the element of $\overline{\mathcal{Y}}_i$ containing t_i, and $\overline{\mathcal{Y}}_i$ is the smallest set containing $\mathcal{Y}_i$ which is closed under intersection. Define $\text{Mini}\, Y_i(t_i) \equiv \bigcap_{y_i\in Y_i(t_i)} y_i$ and let $\text{Mini}\, Y(t) = (\text{Mini}\, Y_i(t_i))_{i\in N}$.

Let $R = (R_i)_{i\in N}$ be an arbitrary profile of reporting correspondences. With any such profile we can associate a unique certifiability configuration $Y^R = (Y_i^R)_{i\in N}$, where $Y_i^R(t_i) \equiv \{R_i^{-1}(r_i) : r_i \in R_i(t_i)\}$ for all $t_i \in T_i$, $i \in N$, and $R_i^{-1}(r_i) \equiv \{t_i \in T_i : r_i \in R_i(t_i)\}$ is the set of types of player i who can send the report r_i. Hence, $\mathcal{Y}_i^R \equiv \{Y_i^R(t_i) : t_i \in T_i\} = \{R_i^{-1}(r_i) : r_i \in \mathcal{R}_i\}$ for all $i \in N$. It is worth mentioning that many different profiles of reporting correspondences can generate the same certifiability configuration.

To characterize in a tractable way the set of all R–certification equilibrium outcomes of a Bayesian game we will prove a generalized version of the revelation principle by defining appropriate canonical communication systems where the profile of reporting correspondences is written as a certifiability configuration. More precisely, given a certifiability configuration Y and its closure $\overline{Y}$, we define a *canonical $\overline{Y}$–communication system* as a $\overline{Y}$–communication system such that $S = T$, $M = A$, $K = 1$, and ν^0 is degenerated. Hence, in a canonical $\overline{Y}$–communication system there is no initial output, there is only one communication period, a report

[12]The set $\mathcal{Y}_i$ is not assumed to be closed under intersection, union or complementation, even if the closure under intersection often seems natural as will be discussed later.

of each player $i \in N$ of type $t_i \in T_i$ is a certificate concerning his type, $y_i \in \overline{Y}_i(t_i)$, a cheap talk signal is a claim about his type, $s_i \in T_i$, and messages sent by the communication system are (recommended) actions.

3.2 Canonical Certification Equilibria

Definition 3 A *canonical $\overline{Y}$–certification equilibrium* of G is a Nash equilibrium of the extended game G_c obtained by adding a canonical $\overline{Y}$–communication system c to G, and in which every player certifies the smallest event concerning his type, reveals truthfully his type, and follows the recommendation of the mediator.

In other words, in a canonical $\overline{Y}$–certification equilibrium each type $t_i \in T_i$ of every player $i \in N$ sends the report $\operatorname{Mini} Y_i(t_i)$, sends the cheap talk signal t_i, and plays the action recommended by the mediator. Hence, such an equilibrium outcome is simply characterized by a recommendation (transition probability) $\nu^* : \overline{\mathcal{Y}} \times T \to \Delta(A)$ satisfying

$$\begin{aligned} &\sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \sum_{a \in A} \nu^*(a \mid \operatorname{Mini} Y(t), t)\, u_i(a; t) \geq \\ &\sum_{t_{-i} \in T_{-i}} p(t_{-i} \mid t_i) \sum_{a \in A} \nu^*(a \mid (\operatorname{Mini} Y_{-i}(t_{-i}), y_i), (t_{-i}, t_i'))\, u_i(a_{-i}, d_i(a_i); t), \end{aligned} \tag{1}$$

for all $i \in N$, t_i, $t_i' \in T_i$, $y_i \in \overline{Y}_i(t_i)$, and $d_i : A_i \to A_i$. The set of canonical $\overline{Y}$–certification equilibrium outcomes is denoted by $\mathcal{E}^*(\overline{Y})$. According to the following theorem, for any profile of reporting correspondences $R = (R_i)_{i \in N}$, the set $\mathcal{E}^*(\overline{Y}^R)$, where $\overline{Y}^R$ is the closure of the certifiability configuration generated by R, exactly *coincides* with the set of all Nash equilibrium outcomes achievable through all R–communication systems. The intuition of this result is similar to the revelation principle for Bayesian games with non-certifiable information (see, e.g., Myerson, 1994). The notable difference, however, is that by taking the closure $\overline{Y}^R$ of the certifiability configuration Y^R generated by the reporting correspondences R we ensure that every information which can be certified by sending different reports at different periods in the original equilibrium can also be certified in the one-period canonical communication system. In that case, the mediator can simulate an outcome equivalent canonical equilibrium and, since any relevant deviation in the canonical equilibrium was already available in the original equilibrium and since only the minimal information is revealed by the mediator through its recommendations, the constructed equilibrium is incentive compatible. The formal proof is detailed in the appendix.

Theorem 1 *The set of R–certification equilibrium outcomes coincides with the set of canonical $\overline{Y}^R$–certification equilibrium outcomes. That is, $\mathcal{E}(R) = \mathcal{E}^*(\overline{Y}^R)$ for all profiles of reporting correspondences R.*

In the following example we give an illustration of the canonical representation and show how the revelation principle applies. We also show that communication equilibria can differ from certification equilibria even if we consider certifiability configurations Y^R allowing players to remain silent, i.e., such that $T_i \in Y_i^R(t_i)$ for all $t_i \in T_i$ and $i \in N$.[13]

Example 2 Consider the game of Figure 2, where $N = \{1, 2\}$, $T = T_1 = \{t^1, t^2, t^3\}$, $A = A_2 = \{a^1, a^2\}$, and consider the following reporting correspondence: $R(t^1) = \{r, r'\}$ and $R(t^2) = R(t^3) = \{r, r', r''\}$. A naive application of the standard revelation principle in this game leads to the conclusion that the complete information outcome $(a^1 \mid t^1, a^2 \mid t^2, a^2 \mid t^3)$ is not implementable since if each type sends a different report to the mediator, then the sender of type t^1 deviates by sending the same report as type t^2 or t^3. Consider on the contrary the canonical representation presented before. The certifiability configuration generated by R is $\mathcal{Y}^R = \{\{t^2, t^3\}, T\}$, so $\overline{Y}^R = Y^R$, $\text{Mini}\, Y^R(t^1) = T$ and $\text{Mini}\, Y^R(t^2) = \text{Mini}\, Y^R(t^3) = \{t^2, t^3\}$. The complete information outcome can be truthfully implemented with the recommendation $\nu^* : \mathcal{Y}^R \times T \to \Delta(A)$ satisfying $\nu^*(a^2 \mid (\{t^2, t^3\}, t^2)) = \nu^*(a^2 \mid (\{t^2, t^3\}, t^3)) = 1$ and $\nu^*(a^1 \mid (y, t)) = 1$ for all other inputs $(y, t) \in \mathcal{Y}^R \times T$. Of course, this outcome is not a communication equilibrium outcome since type t^1 will claim that his type is t^2 or t^3.

	a^1	a^2
t^1	$(0, 1)$	$(1, 0)$
t^2	$(0, 0)$	$(1, 1)$
t^3	$(0, 0)$	$(1, 1)$

Figure 2: Bayesian Game of Example 2.

3.3 One-Period Communication Systems

In this subsection we give a sufficient condition on the profile of reporting correspondences R such that the set of all Nash equilibrium outcomes that can be achieved with all *one-period* R–communication systems coincides with the set of R–certification equilibrium outcomes. The motivation for the restriction to Bayesian games extended with only one-period communication systems is that in some applications one may be interested by the set of equilibria that can be achieved when players are restricted to present only one or few arguments, as it is the case, e.g., in Glazer and Rubinstein's (2001) analysis of debates.

An other interesting example is the configuration examined by Alger and Ma (2003) and Alger and Renault (2002). There, the informed player can be of two different payoff-relevant types, t^1 and t^2, and in addition he can be honest or (possibly) dishonest. The honest player can only reveal his true payoff-relevant

[13] Note that this condition is equivalent to $\bigcap_{t_i \in T_i} R_i(t_i) \neq \emptyset$ for all $i \in N$. In other words, each player can send an uninformative report (i.e., a report which is available whatever his type).

type, whereas the dishonest player can also lie. Denote by t_h^l (t_d^l, resp.) the honest player (dishonest player, resp.) whose payoff-relevant type is t^l, for $l = 1, 2$. The reporting correspondence of the player is thus characterized by $R(t_h^1) = \{t^1\}$, $R(t_h^2) = \{t^2\}$, and $R(t_d^1) = R(t_d^2) = \{t^1, t^2\}$. This correspondence generates the certifiability configuration $\mathcal{Y}^R = \{\{t_h^1, t_d^1, t_d^2\}, \{t_h^2, t_d^1, t_d^2\}\}$, and its closure is $\overline{\mathcal{Y}}^R = \{\{t_h^1, t_d^1, t_d^2\}, \{t_h^2, t_d^1, t_d^2\}, \{t_d^1, t_d^2\}\}$. Consider now the game of Figure 3 with a flat prior probability distribution. It is easy to see that there is a R–certification equilibrium generating the outcome $\mu(a^1 \mid t_d^1) = \mu(a^1 \mid t_d^2) = \mu(a^2 \mid t_h^1) = \mu(a^3 \mid t_h^2) = 1$. However, this equilibrium outcome cannot be achieved with any *one-period* R–communication system since one of the honest type will always imitate the input used by one of the dishonest type. Once multiple communication periods are allowed, a dishonest type can prove to be dishonest by sending two "contradicting" reports (t^1 *and* t^2). This possibility is implicitly introduced by taking the closure of the original certifiability configuration, but is probably not satisfactory given the psychological considerations that motivate the example.[14]

In the following lines we show that if each player is able to certify the intersection of all certifiable events concerning his true type, then considering multiple periods or only single period communication systems is equivalent. Otherwise, as in the previous example, we are not able to provide a simple representation of the set of one-period certification equilibria since different inputs should be used to achieve different possible outcomes, and an initial lottery is thus necessary to ensure the convexity of the set of equilibrium outcomes.

	a^1	a^2	a^3
t_d^1	$(2,1)$	$(0,2)$	$(1,-2)$
t_h^1	$(2,1)$	$(0,2)$	$(1,-2)$
t_d^2	$(2,1)$	$(1,-2)$	$(0,2)$
t_h^2	$(2,1)$	$(1,-2)$	$(0,2)$

Figure 3: One-Period vs Multiple-Period Certification Equilibria.

Definition 4 A certifiability configuration $Y = (Y_i)_{i \in N}$, or an associated profile of reporting correspondences R such that $Y^R = Y$, satisfies the *Minimal Closure Condition* (MCC) if $\text{Mini}\, Y_i(t_i) \in Y_i(t_i)$ for all $i \in N$ and $t_i \in T_i$.

Obviously, a sufficient but not necessary condition for MCC to be satisfied is that each collection of events $\mathcal{Y}_i$ is closed under intersection, i.e., $Y = \overline{Y}$. An other

[14]In particular, following Alger and Renault's (2002) terminology, the "truth-telling honesty" configuration in which an honest player can neither imitate a dishonest player by lying about his payoff-relevant type nor by lying about his ethics becomes equivalent to the previous "equity-minded honesty" configuration in which an honest player is only required to tell the truth concerning his payoff-relevant type.

sufficient condition for a certifiability configuration to satisfy MCC is that it is generated by a profile of reporting correspondences satisfying Green and Laffont's (1986) Nested Range Condition (NRC). More precisely, a profile of reporting correspondences R such that $t_i \in R_i(t_i) \subseteq T_i$ for all $i \in N$ and $t_i \in T_i$ satisfies NRC if for all $i \in N$ and t_i, $t_i' \in T_i$ we have $t_i' \in R_i(t_i) \Rightarrow R_i(t_i') \subseteq R_i(t_i)$. It is not difficult to prove that under NRC the generated certifiability configuration satisfies MCC. However, the converse is not true. Indeed, consider a reporting correspondence as in Example 2: $T = \{t^1, t^2, t^3\}$, $R(t^1) = \{t^1, t^2\}$, $R(t^2) = R(t^3) = T$. NRC is not satisfied since $t^2 \in R(t^1)$ but $R(t^2) \nsubseteq R(t^1)$. However, MCC is satisfied since the generated set of certifiable events, $\mathcal{Y}^R = \{\{t^2, t^3\}, T\}$, is closed under intersection.

Theorem 2 *If R satisfies the minimal closure condition, then the set of one-period R–certification equilibrium outcomes coincides with the set of R–certification equilibrium outcomes.*

An immediate corollary of Theorems 1 and 2 is that under MCC the set of all one-period R–certification equilibrium outcomes exactly coincides with the set of canonical $\overline{Y}^R$–certification equilibrium outcomes. It is also worth mentioning that under MCC a canonical Y^R–certification equilibrium is well defined and that $\mathcal{E}^*(\overline{Y}^R) = \mathcal{E}^*(Y^R)$ even if $\overline{Y}^R \neq Y^R$.

3.4 An Alternative Representation

In this subsection, following the approach of Forges et al. (2002),[15] we present an alternative representation theorem for Bayesian games with certifiable information by constructing, from any given R–communication system, a R^*–communication system in which the set of available inputs of each type t_i of every player i is restricted to a subset $R_i^*(t_i)$ of his set of types (i.e., $R_i^*(t_i) \subseteq T_i$ for all $t_i \in T_i$ and $i \in N$). Such a communication system can be (uniquely) defined for any R–communication system, and the associated set of equilibrium outcomes contains all R–certification equilibrium outcomes. However, in general, it is not equivalent to the set of R–certification equilibrium outcomes because it may contain more outcomes that can actually be achieved with R–communication systems. Nevertheless, natural sufficient conditions are provided for the equivalence to hold.

More precisely, given any profile of reporting correspondences R, let $R_i^*(t_i) \equiv \{s_i \in T_i : \text{Mini}\, Y_i^R(s_i) \in \overline{Y}_i^R(t_i)\}$ for all $t_i \in T_i$ and $i \in N$. That is, in a R^*–communication system the set of all type-dependent inputs that the mediator can receive from each player is a claim concerning his type, where it is implicitly assumed that when some type t_i is reported by player i he also sends the associated certificate $\text{Mini}\, Y_i^R(t_i)$. It is not difficult to check that the profile of correspondences R^* generates the certifiability configuration $\widetilde{\mathcal{Y}}^R = (\widetilde{\mathcal{Y}}_i^R)_{i \in N}$, where for all $i \in N$,

[15] Forges et al. (2002) considered an exchange economy model in which traders are required to present their initial endowments to a market mechanism designer.

$\widetilde{\mathcal{Y}}_i^R \equiv \{\text{Mini}\, Y_i^R(t_i) : t_i \in T_i\}$.[16] Hence, from Theorem 1 we know that the set of R^*–certification equilibrium outcomes coincides with the set of $\widetilde{Y}^R$–certification equilibrium outcomes. Moreover, since players have less possible deviations in a (canonical) $\widetilde{Y}^R$–certification equilibrium than in a (canonical) $\overline{Y}^R$–certification equilibrium, the set of R–certification equilibrium outcomes is included in the set of R^*–certification equilibrium outcomes. The next theorem shows that we can even consider one-period R^*–certification equilibria without initial outputs and without cheap talk signals, where every player truthfully reveals his type and follows the recommendation of the mediator.

Theorem 3 *Every R–certification equilibrium is outcome-equivalent to a one-period R^*–certification equilibrium in which the communication system has no initial output, S is a singleton, $M = A$, and $R_i^*(t_i) \equiv \{s_i \in T_i : \text{Mini}\, Y_i^R(s_i) \in \overline{Y}_i^R(t_i)\}$ for all $i \in N$ and $t_i \in T_i$, and in which strategies are truthful and obedient.*

For example, the complete information outcome obtained in Example 2 can be truthfully implemented with this alternative representation, which gives $R^*(t^1) = \{t^1\}$ and $R^*(t^2) = R^*(t^3) = \{t^1, t^2, t^3\}$. In this example the modification of the reporting correspondence R is irrelevant since the closure of the generated certifiability configuration is not modified ($\overline{Y}^R = \widetilde{Y}^R$). However, in general, the closure of the certifiability configuration generated by R is different from the certifiability configuration generated by R^*, so the inclusion in Theorem 3 may be strict (see Example 1).

The equivalence is restored, for example, if the mediator is able to impose a penalty to any player whose report does not correspond to any equilibrium report, i.e., if for all $i \in N$ and $t_{-i} \in T_{-i}$ there exists $a_{-i} \in A_{-i}$ such that $u_i(a_i, a_{-i}; t) \leq u_i(a'; t)$ for all $a_i \in A_i$, $a' \in A$ and $t_i \in T_i$. This assumption is for instance satisfied in the standard mechanism design framework with transferable utility, where there are $n-1$ agents (with no decision to make) and one uninformed player (the principal) who can make monetary transfers between agents. Alternatively, a mechanism designer or a mediator may be able to directly restrict the set of reporting choices of the individuals (albeit not being able to prevent them to lie), as it is the case when positive disclosures are mandatory. Under one of these conditions, an interesting corollary of Theorem 2 is that under MCC the set of all one-period R–certification equilibrium outcomes exactly coincides with the set of truthful and obedient one-period R^*–certification equilibrium outcomes. This characterization may be very useful in many applications since a truthful and obedient one-period R^*–certification

[16] Of course, when certification possibilities are partial, this implies that players can still lie concerning their true type. For example, if $\text{Mini}\, Y_i^R(s_i) \in \widetilde{Y}_i^R(t_i)$ for $s_i \neq t_i$, then type t_i can certify $\text{Mini}\, Y_i^R(s_i) \neq \text{Mini}\, Y_i^R(t_i)$. This cannot happen, however, if all types can be fully certified, i.e., if $\{t_i\} \in \overline{\mathcal{Y}}_i^R$ for all $i \in N$ and $t_i \in T_i$.

equilibrium is simply characterized by an outcome function $\mu : T \to \Delta(A)$ satisfying

$$\begin{aligned}&\sum_{t_{-i}\in T_{-i}} p(t_{-i} \mid t_i) \sum_{a\in A} \mu(a \mid t)\, u_i(a;t) \geq \\ &\sum_{t_{-i}\in T_{-i}} p(t_{-i} \mid t_i) \sum_{a\in A} \mu(a \mid t_{-i}, t_i')\, u_i(a_{-i}, d_i(a_i); t),\end{aligned} \tag{2}$$

for all $i \in N$, $t_i \in T_i$, $t_i' \in R^*(t_i)$, and $d_i : A_i \to A_i$.

Finally, it is interesting to remark that the approach proposed here allows to make a direct link with Green and Laffont's (1986) framework. Indeed, it can be checked that a profile of reporting correspondences R satisfies NRC if and only if $R = R^*$. As a consequence, if one of the conditions discussed in the previous paragraph is satisfied, then for any profile R we can construct unambiguously, and without loss of generality, another profile R^* satisfying NRC. Otherwise, in the general case, the canonical construction of the representation theorem 1 or 2 should be used.

4 Concluding Remarks

In this paper we have characterized in a tractable way the set of all Nash equilibrium outcomes that can be achieved in Bayesian games in which players have the ability to voluntarily certify and exchange their information through general communication systems. In particular, our framework and results encompass the representation theorem for communication equilibria, as well as existing versions of the revelation principle for principal-agent problems where the set of reports available to the agent is type-dependent.

Since we have considered general communication systems the question of how certification equilibrium outcomes can be implemented in an equilibrium by adding only *unmediated* communication systems to the original Bayesian game was not addressed in this paper and remains the topic of future research. In particular, it should be interesting to investigate whether certification equilibrium outcomes can be implemented with direct communication systems by considering a sufficient number of players (as, e.g., in Bárány, 1992, Ben-Porath, 2003, Forges, 1990, and Gerardi, 2003), by allowing codified messages and bounded computational abilities (as in Urbano and Vila, 2002), or by considering the correlated equilibrium instead of the Nash equilibrium as a solution concept (as in Forges, 1988). It should also be helpful to provide a geometric characterization of the set of Nash equilibrium outcomes achievable with direct communication with certifiable information in two-player games with incomplete information on one side, as is provided by Aumann and Hart (2003) for cheap talk communication. There, the set of communication equilibrium outcomes gives an upper bound for the set of Nash equilibrium outcomes achievable with unmediated communication systems when information is not certifiable. The set of certification equilibrium outcomes characterized in this paper gives exactly the analog of this upper bound for direct communication games with partially verifiable types.

Appendix

To prove the theorems we introduce some lemmas and some additional notations. Denote by $\mathcal{E}(R \mid K=1)$ the set of one-period R–certification equilibrium outcomes, and denote by $\mathcal{E}^{\#}(R^*)$ the set of one-period R^*–certification equilibrium outcomes in which the communication system has no initial output, S is a singleton, $M=A$, and in which strategies are truthful and obedient. Let $Q^R=(Q^R)_{i\in N}$ be the profile of correspondences defined by $Q_i^R(t_i) \equiv \{q_i \in 2^{\mathcal{R}_i} : q_i \subseteq R_i(t_i)\}$ for all $t_i \in T_i$ and $i \in N$. Clearly, we have $Y^{Q^R} = \overline{Y}^{Q^R} = \overline{Y}^R$.

Lemma 1 *If* $\text{Mini}\, Y_i^R(t_i') \in \overline{Y}_i^R(t_i)$*, then* $R_i(t_i') \subseteq R_i(t_i)$*.*

Proof. We show that $R_i(t_i') \nsubseteq R_i(t_i) \Rightarrow \text{Mini}\, Y_i^R(t_i') \notin \overline{Y}_i^R(t_i)$. Let $r_i \in R_i(t_i')$, $r_i \notin R_i(t_i)$. We have $r_i \in R_i(t_i') \Rightarrow R_i^{-1}(r_i) \in \overline{Y}_i^R(t_i') \Rightarrow \text{Mini}\, Y_i^R(t_i') \subseteq R_i^{-1}(r_i)$, and $r_i \notin R_i(t_i) \Rightarrow t_i \notin R_i^{-1}(r_i)$. Thus, $t_i \notin \text{Mini}\, Y_i^R(t_i')$, which implies that $\text{Mini}\, Y_i^R(t_i') \notin \overline{Y}_i^R(t_i)$ since $t_i \in y_i$ for all $y_i \in \overline{Y}_i^R(t_i)$. □

Lemma 2 *For every profile of reporting correspondences* R*,* $\mathcal{E}^*(\overline{Y}^R) \subseteq \mathcal{E}(R)$*. If* R *satisfies MCC, then* $\mathcal{E}^*(\overline{Y}^R) \subseteq \mathcal{E}(R \mid K=1)$*.*

Proof. Let $\nu^* : \overline{\mathcal{Y}}^R \times T \to \Delta(A)$ be any canonical $\overline{Y}^R$–certification equilibrium. We construct an outcome-equivalent R–certification equilibrium as follows. Let c be a R–communication system satisfying $M=A$, $S=T$, $K > |R_i(t_i)|$ for all $i \in N$ and $t_i \in T_i$, ν^k is degenerated for $k=0,1,\ldots,K-1$. In addition, ν^K only depends on the sequence of reporting profiles $r=(r^1,\ldots,r^K) \in \mathcal{R}^K$ and on the cheap talk signals sent in the last communication period (period K), $s^K=(s_1^K,\ldots,s_n^K) \in T$. More precisely, let

$$\nu^K(m,r,s) = \nu^*([\bigcap_{k\in\{1,\ldots,K\}} R_i^{-1}(r_i^k)]_{i\in N}, s^K),$$

for all $(m,r,s) \in M^K \times \mathcal{R}^K \times T^K$. Since $\overline{\mathcal{Y}}_i^R = \{\bigcap_{k\in\{1,\ldots,K\}} R_i^{-1}(r_i^k) : r_i \in \mathcal{R}_i^K\}$ and $\bigcap_{r_i \in R_i(t_i)} R_i^{-1}(r_i) = \text{Mini}\, Y_i^R(t_i)$ for all $i \in N$, the strategy which consists for each type t_i of every player i in sending every report in $R_i(t_i)$ during the communication phase, revealing his true type in the last communication period and following the recommendation of the mediator is, by the definition of the original canonical $\overline{Y}^R$–certification equilibrium and the construction of c, a Nash equilibrium of G_c. This equilibrium is clearly outcome-equivalent to ν^*. Similarly, to prove the second part of the lemma let $\nu(m,r,s) = \nu^*([R_i^{-1}(r_i)]_{i\in N}, s)$ for all $(m,r,s) \in M \times \mathcal{R} \times T$ and remark that under MCC, for all $i \in N$ and $t_i \in T_i$, there exists $r_i \in R_i(t_i)$ such that $R_i^{-1}(r_i) = \text{Mini}\, Y_i^R(t_i)$. $\square$

Lemma 3 *For every profile of reporting correspondences R, $\mathcal{E}(R) \subseteq \mathcal{E}(Q^R \mid K = 1)$.*

Proof. Consider any Nash equilibrium of any communication game G_c where c is a R–communication system. We construct an outcome-equivalent one-period Q–certification equilibrium where $Q = Q^R$, $M = A$, the initial lottery is degenerated, the transition probability is $\pi : (\prod_{i \in N} 2^{\mathcal{R}_i}) \times T \to \Delta(A)$, each player i of type t_i follows the recommendation generated by π, sends the report $R_i(t_i) \in Q(t_i)$ and reveals his true type. That is, $\sigma_i(t_i) = (R_i(t_i), t_i)$ and $\delta_i(a_i \mid a_i, r_i, s_i, t_i) = 1$ for all $t_i \in T_i$, $a_i \in A_i$, $(r_i, s_i) \in Q_i(t_i) \times T_i$ and $i \in N$. If every player i sends an input $(R_i(s_i), s_i)$ for some $s_i \in T_i$, then π simulates the action profile played in the original equilibrium when the type profile is $s = (s_1, \ldots, s_n) \in T$. Clearly, this constructed mechanism generates the original equilibrium outcome. To verify that it is incentive compatible we must verify that for every player i, no type t_i has an incentive to deviate from $(R_i(t_i), t_i)$ to $(q_i, s_i) \neq (R_i(t_i), t_i)$ for all $(q_i, s_i) \in Q_i(t_i) \times T_i$. If $(q_i, s_i) = (R_i(s_i), s_i)$ (unobservable deviation), then $R_i(s_i) \subseteq R_i(t_i)$ (because $R_i(s_i) \in Q_i(t_i) \Rightarrow R_i(s_i) \subseteq R_i(t_i)$), which means that type t_i already had the possibility to imitate type s_i's communication strategy under the original equilibrium. If $(q_i, s_i) \neq (R_i(s_i), s_i)$ (observable deviation), then π simulates the outcome generated by a deviation of player i to, e.g., an unconditional sequence of K reports of any single report in q_i and K cheap talk signals in S_i under the original equilibrium. This deviation was already available to type t_i since $q_i \in Q_i(t_i) \Rightarrow q_i \subseteq R_i(t_i)$. $\square$

Lemma 4 *For every profile of reporting correspondences R, $\mathcal{E}(R \mid K = 1) \subseteq \mathcal{E}^*(\overline{Y}^R)$.*

Proof. The proof is similar to the proof of Lemma 3. Consider any Nash equilibrium of any communication game G_c where c is a one-period R–communication system. We construct an outcome-equivalent canonical $\overline{Y}$–certification equilibrium $\nu^* : \overline{\mathcal{Y}} \times T \to \Delta(A)$, where $\overline{Y} = \overline{Y}^R$, as follows. If every player i sends an input $(\text{Mini}\, Y_i(s_i), s_i)$ for some $s_i \in T_i$, then ν^* simulates the action profile played in the original equilibrium when the type profile is $s \in T$. If some player i sends an input $(y_i, s_i) \neq (\text{Mini}\, Y_i(s_i), s_i)$, then ν^* simulates the outcome generated by player i's deviation to some report r_i such that $y_i \subseteq R_i^{-1}(r_i)$ and some cheap talk signal in S_i under the original equilibrium. This deviation was already available to type t_i since $y_i \in \overline{Y}_i(t_i) \Rightarrow t_i \in y_i \subseteq R_i^{-1}(r_i) \Rightarrow r_i \in R_i(t_i)$. It remains to show that type t_i has no incentive to send an input $(\text{Mini}\, Y_i(s_i), s_i)$ for $s_i \neq t_i$. This is obtained by Lemma 1 since $\text{Mini}\, Y_i(s_i) \in Y_i(t_i)$, so $R_i(s_i) \subseteq R_i(t_i)$, which means that an equivalent deviation was already available under the original equilibrium. $\square$

Proof of Theorem 1. Lemma 4 gives $\mathcal{E}(Q^R \mid K = 1) \subseteq \mathcal{E}^*(\overline{Y}^{Q^R}) = \mathcal{E}^*(\overline{Y}^R)$. Therefore, by Lemmas 2 and 3 we get $\mathcal{E}(Q^R \mid K = 1) \subseteq \mathcal{E}^*(\overline{Y}^R) \subseteq \mathcal{E}(R) \subseteq \mathcal{E}(Q^R \mid K = 1)$, so $\mathcal{E}^*(\overline{Y}^R) = \mathcal{E}(R)$. □

Lemma 5 *For every profile of reporting correspondences* R, $\mathcal{E}^{\#}(R^*) = \mathcal{E}^*(\widetilde{Y}^R)$.

Proof. Clearly, we have $\mathcal{E}^{\#}(R^*) \subseteq \mathcal{E}(R^*)$. In addition, Theorem 1 gives $\mathcal{E}(R^*) = \mathcal{E}^*(\widetilde{Y}^R)$ because the profile of reporting correspondences R^* generates the certifiability configuration $\widetilde{Y}^R$. Thus, we have to show that $\mathcal{E}^*(\widetilde{Y}^R) \subseteq \mathcal{E}^{\#}(R^*)$. Let $\nu^* : \widetilde{\mathcal{Y}}^R \times T \to \Delta(A)$ be any canonical $\widetilde{Y}^R$–certification equilibrium. We have to show that $\mu : T \to \Delta(A)$, where $\mu(a \mid t) = \nu^*(a \mid \operatorname{Mini} Y^R(t), t)$, induces a truthful and obedient one-period R^*–certification equilibrium outcome (an outcome in $\mathcal{E}^{\#}(R^*)$), i.e., that Equation (2) on page 15 is satisfied for all $t'_i \in R^*(t_i)$ and $d_i : A_i \to A_i$. Since $t'_i \in R^*(t_i) \Leftrightarrow \operatorname{Mini} Y_i^R(t'_i) \in \widetilde{Y}_i^R(t_i)$, this condition is implied by the fact that ν^* is a $\widetilde{Y}^R$–certification equilibrium outcome (see Equation (1) on page 10 with $Y = \overline{Y} = \widetilde{Y}^R$). □

Proof of Theorem 2. Let R be a profile of reporting correspondences satisfying MCC. By Lemma 4 we have $\mathcal{E}(R \mid K = 1) \subseteq \mathcal{E}^*(\overline{Y}^R)$, and by Lemma 2 we have $\mathcal{E}^*(\overline{Y}^R) \subseteq \mathcal{E}(R \mid K = 1)$, so $\mathcal{E}^*(\overline{Y}^R) = \mathcal{E}(R \mid K = 1)$. Thus, by Theorem 1 we get $\mathcal{E}(R) = \mathcal{E}(R \mid K = 1)$. □

Proof of Theorem 3. We have $\mathcal{E}(R) = \mathcal{E}^*(\overline{Y}^R)$ by Theorem 1, $\mathcal{E}^*(\overline{Y}^R) \subseteq \mathcal{E}^*(\widetilde{Y}^R)$ because players have less possible deviations in a canonical $\widetilde{Y}^R$–certification equilibrium than in a canonical $\overline{Y}^R$–certification equilibrium, and $\mathcal{E}^*(\widetilde{Y}^R) = \mathcal{E}^{\#}(R^*)$ by Lemma 5. Consequently, $\mathcal{E}(R) \subseteq \mathcal{E}^{\#}(R^*)$. □

References

Alger, I. and C.-t. A. Ma (2003): "Moral Hazard, Insurance, and some Collusion," *Journal of Economic Behavior and Organization*, 50, 225–247.

Alger, I. and R. Renault (2002): "Screening Ethics when Honest Agents Care about Fairness," mimeo.

Aumann, R. J. (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1, 67–96.

Aumann, R. J. and S. Hart (2003): "Long Cheap Talk," *Econometrica*, forthcoming.

Baliga, S. and S. Morris (2002): "Co-ordination, Spillovers, and Cheap Talk," *Journal of Economic Theory*, 105, 450–468.

Bárány, I. (1992): "Fair Distribution Protocols or How the Players Replace Fortune," *Mathematics of Operations Research*, 17, 327–340.

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

Ben-Porath, E. (2003): "Cheap Talk in Games with Incomplete Information," *Journal of Economic Theory*, forthcoming.

Bull, J. and J. Watson (2002): "Hard Evidence and Mechanism Design," mimeo, University of California, San Diego.

Che, Y.-K. and I. Gale (2000): "The Optimal Mechanism for Selling to a Budget-Constrained Buyer," *Journal of Economic Theory*, 92, 198–233.

Crawford, V. P. and J. Sobel (1982): "Strategic Information Transmission," *Econometrica*, 50, 1431–1451.

d'Aspremont, C., S. Bhattacharya, and L.-A. Gérard-Varet (2000): "Bargaining and Sharing Innovative Knowledge," *Review of Economic Studies*, 67, 255–271.

Deneckere, R. and S. Severinov (2001): "Mechanism Design and Communication Costs," mimeo, University of Wisconsin-Madison.

Dye, R. A. (2001): "An Evaluation of "Essays on Disclosure" and the Disclosure Literature in Accounting," *Journal of Accounting and Economics*, 32, 181–235.

Farrell, J. and M. Rabin (1996): "Cheap Talk," *Journal of Economic Perspectives*, 10, 103–118.

Forges, F. (1986): "An Approach to Communication Equilibria," *Econometrica*, 54, 1375–1385.

——— (1988): "Can Sunspots Replace a Mediator?" *Journal of Mathematical Economics*, 17, 347–368.

——— (1990): "Universal Mechanisms," *Econometrica*, 58, 1341–1364.

——— (1993): "Five Legitimate Definitions of Correlated Equilibrium in Games with Incomplete Information," *Theory and Decision*, 35, 277–310.

Forges, F., J.-F. Mertens, and R. Vohra (2002): "The Ex Ante Incentive Compatible Core in the Absence of Wealth Effects," *Econometrica*, 70, 1865–1892.

Gerardi, D. (2003): "Unmediated Communication in Games with Complete and Incomplete Information," *Journal of Economic Theory*, forthcoming.

Glazer, J. and A. Rubinstein (2001): "Debates and Decisions: On a Rationale of Argumentation Rules," *Games and Economic Behavior*, 36, 158–173.

Green, J. R. and J.-J. Laffont (1986): "Partially Verifiable Information and Mechanism Design," *Review of Economic Studies*, 53, 447–456.

Grossman, S. J. (1981): "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 24, 461–483.

Grossman, S. J. and O. D. Hart (1980): "Disclosure Laws and Takeover Bids," *Journal of Finance*, 35, 323–334.

Koessler, F. (2002): "Strategic Knowledge Sharing in Bayesian Games," mimeo, THEMA, Université de Cergy-Pontoise.

——— (2003): "Persuasion Games with Higher-Order Uncertainty," *Journal of Economic Theory*, forthcoming.

Krishna, V. and J. Morgan (2002): "The Art of Conversation," mimeo.

Lipman, B. L. and D. Seppi (1995): "Robust Inference in Communication Games with Partial Provability," *Journal of Economic Theory*, 66, 370–405.

Milgrom, P. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380–391.

Myerson, R. B. (1982): "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems," *Journal of Mathematical Economics*, 10, 67–81.

——— (1986): "Multistage Games with Communication," *Econometrica*, 54, 323–358.

——— (1994): "Communication, Correlated Equilibria and Incentive Compatibility," in *Handbook of Game Theory*, ed. by R. J. Aumann and S. Hart, Elsevier Science B. V., vol. 2, chap. 24, 827–847.

Okuno-Fujiwara, A., M. Postlewaite, and K. Suzumura (1990): "Strategic Information Revelation," *Review of Economic Studies*, 57, 25–47.

Seidmann, D. J. and E. Winter (1997): "Strategic Information Transmission with Verifiable Messages," *Econometrica*, 65, 163–169.

Shin, H. S. (1994): "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, 64, 253–264.

——— (1998): "Adversarial and Inquisitorial Procedures in Arbitration," *Rand Journal of economics*, 29, 378–405.

——— (2003): "Disclosures and Asset Returns," *Econometrica*, 71, 105–133.

Urbano, A. and J. E. Vila (2002): "Computational Complexity and Communication: Coordination in Two-Player Games," *Econometrica*, 70, 1893–1927.

Verrecchia, R. E. (2001): "Essays on Disclosure," *Journal of Accounting and Economics*, 32, 97–180.

Wolinsky, A. (2002): "Eliciting Information from Multiple Experts," *Games and Economic Behavior*, 41, 141–160.

——— (2003): "Information Transmission when the Sender's Preferences are Uncertain," *Games and Economic Behavior*, forthcoming.