When and Why? A Critical Survey on Coordination Failure in the Laboratory

GIOVANNA DEVETAG

Department of Law and Management, University of Perugia

Via Pascoli, 20, 06123 Perugia, Italy

tel.: (39) 075 5855271, fax: (39) 075 5855257

e-mail: devetag@unipg.it, mgiovanna.devetag@libero.it

ANDREAS ORTMANN

CERGE-EI, Charles University and Academy of Sciences of the Czech Republic

P.O.BOX 882, Politickych veznu 7, 111 21 Prague, Czech Republic

tel.: (420) 242 30 117, fax: (420) 242 11 374, 242 27 143

e-mail: aortmann@cerge-ei.cz, aortmann@yahoo.com

Journal of Economic Literature Classification Numbers: C72, C92

Key Words: coordination games, Pareto-ranked equilibria, payoff-asymmetric equilibria, stag-hunt games, optimization incentives, robustness, coordination, coordination failure

Abstract

Coordination games with Pareto-ranked equilibria have attracted major attention over the past two decades. Two early path-breaking sets of experimental studies were widely interpreted as suggesting that coordination failure is a common phenomenon in the laboratory. We identify the major determinants that seem to affect the incidence, and/or emergence, of coordination failure in

the lab and review critically the existing experimental studies on coordination games with Pareto-ranked equilibria since that early evidence emerged. We conclude that there are many ways to engineer coordination successes.

1. Introduction

> "Several basic conclusions have emerged from this research: Coordination failure is common …" (Camerer, 2003, p. 403)

Coordination games with Pareto-ranked equilibria, or "payoff-asymmetric" coordination games (Camerer, 2003, section 7.4), have attracted major attention over the past two decades. Two path-breaking and frequently cited early sets of experimental studies (namely, Van Huyck, Battalio and Beil [VHBB henceforth], 1990, 1991, and Cooper, DeJong, Forsythe and Ross [CDFR], 1990, 1992) have been interpreted as suggesting that coordination failure is a common phenomenon in the laboratory. Coordination failure describes either of these events: Failure to coordinate on any one of the multiple equilibria or failure to coordinate on the payoff dominant equilibrium. The latter meaning has been used by VHBB (1990, 1991) who pointed out that this meaning was the convention that was developing then in the literature on macroeconomic coordination games. We follow that convention below.

The claim that coordination failure might be a common phenomenon prompted a steady flow of robustness tests. In this article we review critically order-statistic games like those in VHBB (1990) and VHBB (1991) and stag-hunt games like those in CDFR

(1992). We are well aware that these labels are somewhat misleading, as both are coordination games with Pareto-ranked equilibria. We stick to these labels mainly for historic reasons – the two sets of experimental studies that initiated the experimental literature on coordination games with Pareto-ranked equilibria, and the rather different experimental paradigms used to implement them.

Our research strategy consists of a qualitative review (NOTE1) of the available evidence in an attempt to classify the major classes of structural, cognitive, and behavioral determinants that seem to affect coordination failure in the lab (NOTE2). We also reflect briefly on the external validity of the currently available set of laboratory coordination game studies.

2. The class of games under consideration

Order-statistic games. The payoff function of a generic order-statistic game can be represented as follows:

(1) $\Pi_i = f(OS - |e_i – OS|)$

where OS stands for the order statistic (e.g., the median or the minimum, or something else), e_i denotes player i's choice of an ordered set of numbers which is meant to represent an ordered set of efforts (NOTE3), $|e_i – OS|$ denotes the (symmetric) deviation cost, and f is some scalar function of these terms. Obviously, the terms can be arbitrarily modified by setting the coefficients of the two terms on the RHS not equal to 1, or by squaring the second term, or by defining the deviation costs asymmetrically, etc.

VHBB (1990, 1991) used the following earnings tables for their two path-breaking studies.

- Table 1 here -

Table 1: Earnings table for the "Median game" (Table Γ in VHBB, 1991)

		Median value of X chosen						
		7	6	5	4	3	2	1
Your choice of X	7	1.30	1.15	0.90	0.55	0.10	-0.45	-1.10
	6	1.25	1.20	1.05	0.8	0.45	0.00	-0.55
	5	1.10	1.15	1.10	0.95	0.70	0.35	-0.10
	4	0.85	1.00	1.05	1.00	0.85	0.60	0.25
	3	0.50	0.75	0.90	0.95	0.90	0.75	0.50
	2	0.05	0.40	0.65	0.80	0.85	0.80	0.65
	1	-0.5	-0.05	0.3	0.55	0.70	0.75	0.70

- Table 2 here -

Table 2: Earnings table for the "Minimum game" (Table A in VHBB, 1990)

		Smallest value of X chosen						
		7	6	5	4	3	2	1
Your choice of X	7	1.30	1.10	0.90	0.70	0.50	0.30	0.10
	6	-	1.20	1.00	0.80	0.60	0.40	0.20
	5	-	-	1.10	0.90	0.70	0.50	0.30
	4	-	-	-	1.00	0.80	0.60	0.40
	3	-	-	-	-	0.90	0.70	0.50
	2	-	-	-	-	-	0.80	0.60
	1	-	-	-	-	-	-	0.70

Note that the payoff-dominant, or efficient, equilibrium is in the upper left corner for both the Minimum game (Table 2) and the Median game (Table 1) while the secure action induces an equilibrium (the secure equilibrium from here on) in the lower right corner for the Minimum game and two rows up from the bottom in the Median game. Both games feature seven (identical) Pareto-ranked pure-strategy equilibria on the main diagonal. There is a tension between the secure action – the lowest action in the Minimum game, and the third lowest in the Median game - and the action required for the efficient equilibrium.

Importantly, the payoffs in the triangular area above the main-diagonal are not the same: For the Minimum game deviation costs are linear and feature positive payoffs only, whereas for the Median game they are highly non-linear, leading to negative payoffs in the upper right corner and lower left corner. This nonlinearity (and the negative payoffs that it induces) counteracts, and possibly neutralizes, the more forgiving Median statistic. These differences in deviation costs confound the comparison between Median and Minimum game experiments. In fact, the labeling of the games is unfortunate because it distracts from the consequences of the different parameterizations.(NOTE4)

Stag-hunt games. This class of games, like order-statistic games, feature multiple (in this case, typically two) pure-strategy equilibria that are Pareto-ranked. Payoffs result from the strategic interaction of two players with two action choices each. CDFR (1992) contained the paradigmatic example of this class of games, sg(1,x,y,z) = 1,000g(1,0,0.8,0.8), where g is normalized to 1, s is a scalar function here taking on the value 1000, $x<z$, $y<1$, and $x,y,z \in [0,1)$:

		Other player's choice	
		2	1
Your	2	1,000	0
Choice	1	800	800

Like the order-statistic games discussed earlier, the payoff-dominant equilibrium is in the upper left corner while the secure equilibrium is in the lower right corner: There is thus a tension between the risky action (required for the efficient equilibrium) and the secure action. The concept of security will always select Choice 1 and therefore, quite possibly, select secure but unattractive equilibria. A more persuasive solution concept is risk dominance (NOTE5) which, for some values of y and z, selects the efficient equilibrium. Essentially, this is the case when the secure action choice is not attractive enough.

Classes of determinants of coordination outcomes. Prominent structural determinants of coordination failure are thus the *attractiveness of the secure, or maximin, strategy and the riskiness of the other action choices*, which are partially defined by the type and strength of deviation costs, as well as the *coordination requirements determined by order statistic* and *group size*, and the opportunities for *shared experience, interaction, and informational feedback*. These structural factors may be usefully labeled exogenous risk characteristics because they are fully under the control of the experimenter.

Cognitive and behavioral determinants – e.g., subjects' understanding of the payoff matrices, or the effects that potentially *negative payoffs* might have on subjects -- are those not fully under the control of the experimenter. In light of the well-documented sensitivity of outcomes to initial conditions, discussed below, in some games (e.g., the "Median" games in VHBB, 1991) these issues seem of obvious importance as they add to the exogenous risk characteristics endogenous ones that VHBB (1990, 1991) called "strategic uncertainty."

We now turn our attention to what we know empirically about these determinants.

3. Laboratory evidence of coordination failures and successes

3.1. *Attractiveness of the secure strategy and riskiness of the other action(s)*

Order statistic games. Was efficiency psychologically salient in VHBB (1990, 1991) or were competing concepts such as security, or risk dominance, more salient?

The key result of VHBB (1990) is the stable and speedy unraveling of action choices to the worst of the seven Pareto-ranked strict equilibria. Between 14 – 16 participants played the stage game repeatedly receiving only information about the minimum after each stage. Several other experimenters replicated this unraveling result with the same payoff matrix but the number of subjects varying from 6 – 14 (e.g., Cachon and Camerer, 1996; Bornstein et al., 2002; Blume and Ortmann, 2007; Chaudhuri et al., 2005), as well as with structurally similar payoff matrices with slightly more or less action choices (e.g., Berninghaus and Ehrhart, 1998; Knez and Camerer, 1994; Weber et al., 2001; Brandts and Cooper, 2004, 2006).

The key result of VHBB (1991) is the influential role that the initial action choices play as the initial median constituted a strong precedent from which subjects had trouble extracting themselves. This result has also been replicated in some of the references cited above.

Because of their remarkable results on coordination failure, VHBB (1990, 1991) drew considerable attention and a steady flow of attempts to test their robustness.(NOTE6)

Every choice between a secure and a (set of) riskier actions is ultimately a function of the expected values of the available choices. The higher the expected value of the secure action the more likely it is to undermine the risky actions, and vice versa.

Brandts and Cooper (2004) address this issue head-on. Studying coordination in a minimum effort game with five effort levels, and keeping the payoff associated with the minimum constant, they vary the payoff associated with the efficient equilibrium and observe higher incidence of coordination success as the efficient equilibrium becomes significantly more attractive.

Other authors have explored the robustness of coordination by manipulating the type and strengths of the deviation costs. Berninghaus and Ehrhart (1998) introduced longer time horizons (scaling down the per-round payoffs accordingly) in a minimum game, so as to lower the opportunity cost of exploration. The number of rounds had the hypothesized effect although they did not bring about complete convergence to the Pareto-efficient outcome, and there was no difference in the distribution of initial choices. Goeree and Holt (2005) varied the magnitude of deviation costs (what they call the "effort cost") in series of one-shot minimum and median games with two or three players: they find that the impact of deviation costs on coordination success is significant, although their use of a finer action grid and of continuous strategies, along with the different matching protocol, renders their results not directly comparable with those of VHBB (1990, 1991).

Van Huyck et al. (2001) explore the consequences of a finer action grid (as well as the impact of the order statistic and the number of players). Letting their subjects choose among 101 actions (and letting them run through twice the numbers of rounds), the authors find that local exploration is "skewed in the direction of efficiency." It is possible, and likely – in light of the observed perfect correlation between "creeping up" and time in some of the treatments, and the results by Berninghaus and Ehrhart (1998), that this result is due to both the refined action space and to the increased number of rounds. The refined action space, combined with an extended time horizon, may also have caused a similar drift toward efficiency in Van Huyck et al. (1997).

Stag-hunt games. Both CDFR (1990) and CDFR (1992) were concerned with stag-hunt games of the sg(1,x,y,z) variety where $x < y = z$. CDFR (1990), however, embedded the stag-hunt games into a larger 3 x 3 matrix that featured a Pareto dominant outcome that was induced by a dominated strategy. The key question was whether the Pareto-dominant equilibrium would always be selected. The answer was no as dominated strategies that could have induced the Pareto-dominant equilibria were not selected.

CDFR (1992) also explored whether the results in CDFR (1990) were robust to the use of one-way and two-way communication. Coordination failure turned out to be endemic in the no-communication conditions and to still be significant with one-way communication, but was eliminated with two-way communication. We return to the issue of communication below.

It is important to mention that these coordination failures came about under a matching protocol that differed from the one used by VHBB (1990, 1991) and others. Specifically, while VHBB and others nearly always used a multi-player fixed matching protocol, CDFR (1990, 1992) typically used two-player sequences of one-shot games – a matching scheme known informally as highway protocol in which each player plays every other player exactly once. A related matching scheme is the random matching protocol in which a player is randomly matched with one of the other players, possibly repeatedly. The matching protocol can affect outcomes. (More on this below.)

Schmidt et al. (2003), in an article closely related to Battalio et al. (2001), systematically vary measures of payoff-dominance and risk-dominance and find – both for random matching and fixed matching protocols -- that players react to changes in risk-dominance but not payoff-dominance. This result is at odds with the results in Battalio et al. (2001) and Clark et al. (2001). Importantly, and in contradiction to the message the title of their paper suggests, subjects selected “the payoff dominant strategy more often than not” (p. 298), with this statement applying to almost all treatments regardless of the matching protocol.

Arguably the most intriguing article in this area – because its results seem dramatically at odds with claims that coordination failure is common -- is Rankin, Van Huyck and Battalio [RVHB] (2000). The authors use a scaled-up version of g(1,0,x,x) where x is, for each round, drawn randomly from the unit-interval and then, ever so slightly, perturbed. Taking the cue from Kreps's (1990) argument that experience with precisely the same game in precisely the same situation is hardly a way to instill trust in the generalizability of laboratory results, RVHB had their subjects play a sequence of 75 games, scrambling the action labels so that the payoff dominant equilibrium and the secure equilibrium would not show up in the same cell throughout.. They found a high percentage of efficient play both when $x < 0.5$ (when the payoff dominant and risk dominant equilibrium coincided) and when $x > 0.5$ (when the payoff dominant and the risk dominant equilibrium were at opposite ends of the main diagonal): for the first 10 periods 65% (85%) of choices corresponded to the efficient action when $x > 0.5$ ($x < 0.5$), with about 90% (almost 100%) of the choices corresponded to the efficient action in the last 10 periods. Thus, payoff dominance clearly carried the day. RVHB point out that their set-up inhibits learning from experience and focuses subjects on the exploration of deductive principles. In addition, in about half of the rounds subjects faced a situation in which payoff-dominance and risk-dominance selected the same equilibrium. It probably also helped that subjects were told that "you will remain grouped with the same seven other participants for the next 75 rounds." This formulation is likely to have translated in most subjects' minds into the belief that they were going to see each of the seven other participants about 10 or 11 times, a potential source of trust-building.

3.2. *Order statistic and group size*

The coordination requirement in order statistic games is related both to the particular order statistic used to calculate payoffs and to the group size. Intuition suggests that, all other things constant, in the minimum effort game it is riskier to pick the efficient action in large groups than in small groups.

Van Huyck et al (2001) directly tested the claim by crossing two group sizes (5 and 7) and two order statistics (2 and 4) in a 2 x 2 design that also featured a dramatically increased action space (101 actions) and a relatively large number of periods. Among the many interesting results – contradicting the Berninghaus and Ehrhart (1998) results about initial values – is that the variations in the order statistic and group sizes influenced behavior in the first round, with subjects reacting more strongly to differences in the order statistic than in group size.

3.3. *Shared experience, interaction, and other informational issues*

A precedent results from shared experience (Lewis, 1969) and creates expectations on the part of the participants about what happens next. Precedents are created when players interact repeatedly with the same players, as in VHBB (1991) or the fixed matching treatments of VHBB (1990). Shared experience can also be induced, ex ante, via precedents established in other contexts. The possibility of observing the actions of other players, or the possibility to inform other players of one's intentions through costly or costless pre-play communication is among the other informational issues that affect the outcomes of coordination games.

Order statistic games. VHBB label precedents from other games "weak precedents" to distinguish them from the "strong precedent" established in a previous round of the same game. This terminology is not always descriptive. Weber (2005) has demonstrated that if trust is built slowly and new participants are made aware of the group's history, efficient precedents can spill over from n-person weak-link games to (n+1)-person weak-link games (but see also Knez and Camerer, 2000 who demonstrate that precedents can be fragile expectational assets.)

The effect of information has been studied in a number of experiments. The results are mixed: Berninghaus and Ehrhart (2001) and Brandts and Cooper (2006) suggest post-play information is efficiency-enhancing; Devetag (2005) and the full information treatment in VHBB (1990) suggest no effects.

Other studies investigate the role of costless ("cheap talk") and costly pre-play communication. Van Huyck et al. (1993) -- by auctioning off the right to play – used costly (but tacit) information to overcome coordination failure completely. Turning from costly to costless messages, Blume and Ortmann (2007) test the effect of cheap talk both in the Minimum and Median game. They find that costless messages with minimal information content, when added to games with Pareto-ranked equilibria, can facilitate both quick convergence to, and participants' initial coordination on, the Pareto-dominant equilibrium. Cheap talk is thus a substitute for other efficiency-enhancing characteristics such as a more forgiving order statistic, smaller group size, or step size, or a refined actions space. See also Burton and Sefton (2004) for similar results in a closely related class of games. Chaudhuri et al. (2005) find that cheap talk is efficiency-enhancing in intergenerational minimum effort game experiments, as the quality of advice given is positively related to the probability of coordination success.(NOTE7)

Stag-hunt games. While a number of papers have studied the effect of changes in the payoff matrix, relatively few authors have studied the effect of the kind of design and implementation details studied in order statistic games. This state of affairs is deplorable because these issues may be more important than structural characteristics of the payoff matrix.

Interestingly, the impact of pre-play communication in terms of costly signals has not been studied extensively in the context of stag-hunt games. Aumann (1990) conjectured that the impact of cheap talk would depend significantly on the structure of the payoff matrix. Specifically, in g(1,0,0.9,0.7) messages expressing the intent to shoot for the payoff-dominant equilibrium would not be credible because it is in a player's interest to entice the other player to do so and then cheat. In contrast, in g(1, 0,0.8,0.8) such an expression would not be self-serving. Clark et al. (2001) provide evidence in support of this conjecture when comparing no communication and two-way communication (also see Charness, 2000).

Duffy and Feltovich (2002, 2006) study the impact of words, deeds and lies on behavior in prominent strategic situations, including the stag-hunt game. If cheap talk is credible then words indeed speak louder than deeds. While subjects are quite honest to start with, the possibility of being caught lying improves the already high coordination even more.

Bangun et al. (2006), like Van Huyck et al. (1992), study the effects of external assignments. They found significant effects of external assignments in three-action scenarios with Pareto-ranked equilibria that did not have any tension between the payoff-dominant and risk-dominant outcomes. Bangun et al., (2006), using a game with tension between the two outcomes, found -- in contrast to the results of Chaudhuri et al. (2005) – that recommendations by the experimenter to play the risky strategy induce the efficient equilibrium under both “common knowledge” and “almost common knowledge”.

Among the few papers that have explored implementation issues in the stag-hunt game, Clark and Sefton (2001) investigate the role of interaction structure. Their experiment involves the play of a stag-hunt game and the use both of the highway protocol and the fixed matching protocol. The latter may influence behavior in a variety of ways, the most obvious of which is the possibility to use precedent. However, an additional, more subtle way in which a fixed matching protocol may alter behavior is through the possibility of costly signaling that it offers players. This type of signaling is costly insofar as it implies the possibility of having zero payoff rounds initially. In order to distinguish between the two phenomena, Clark and Sefton investigate first round behavior, in which only the impact of signaling should be observed. Their data show that, indeed, in the first round of play the frequencies of choice of the risky action were 0.3 in the highway protocol and 0.6 in the fixed matching protocol, a significant difference. Moreover, the fixed matching protocol reduced the instances of disequilibrium outcomes and increased the overall proportion of risky choices across rounds.

3.4. *Negative payoffs*

Order-statistic games. Although an affine transformation of payoffs does not change the structure of equilibria in a coordination game, there is some evidence (albeit by no means undisputed, see e.g., List, 2004; Plott and Zeiler, 2005) that framing outcomes as gains or losses is not neutral with respect to behavior. Drawing on VHBB (1991, 1993), Cachon and Camerer (1996) investigate loss avoidance as a selection principle: if people follow loss avoidance, they should avoid playing strategies that result in certain losses if strategies leading to potential gains are available. They find that loss avoidance functions as a selection principle in the median as well as the minimum effort game, inducing coordination on the Pareto-dominant equilibrium. Here, too, no studies exist (yet) that investigate the role of negative payoffs in a systematic way, though it would seem to be called for given the likelihood that the initial choices in the classic Median and Minimum game were at least partially affected by the presence or absence of negative payoffs. A reasonable conjecture would be that the prominent negative payoffs in the upper right, and lower left, corner of the Median game table of VHBB (1991) did affect people's choices, and were responsible for the clustering of initial choices slightly above the secure action.

Stag-hunt games. Rydval and Ortmann (2005), and Feltovich et al. (2005) tested experimentally the Cachon-Camerer conjecture that loss avoidance might also work its magic in stag-hunt games. Both papers suggest that loss avoidance may indeed be a (weak) selection principle in stag-hunt games, especially if losses are certain for a chosen action.

4. Discussion

What we have learned since VHBB (1990, 1991) and CDFR (1990, 1992)(NOTE8):

- Lower attractiveness of the secure action relative to the risky action required for the efficient equilibrium is efficiency-enhancing (e.g., Brandts and Cooper, 2004).
- Low (zero) deviation costs are efficiency enhancing (e.g., VHBB, 1990; Goeree and Holt, 2005; Battalio et al., 2001).
- Lower costs of experimentation such as increasing the number of rounds while keeping the overall earnings roughly the same, or refining the actions space, or some combination thereof, are efficiency-enhancing (e.g., Berninghaus and Ehrhart, 1998; Van Huyck et al., 2001).
- Less stringent coordination requirements (i.e., a smaller group size or a less stringent order statistic) are efficiency-enhancing (e.g., VHBB, 1990; Van Huyck et al., 2001).

- Fixed matching protocols are efficiency enhancing (e.g., VHBB, 1990; Clark and Sefton, 2001; Schmidt et al., 2003).

- Repeated encounters are efficiency enhancing even under random matching schemes if the experimental design and implementation focuses subjects on deductive principles (e.g., RVHB, 2000; see also Schmidt et al., 2003).

- Providing full information feedback seems efficiency enhancing in "small" groups (e.g., Berninghaus and Ehrhart, 2001; Brandts and Cooper, 2006; but see VHBB 1990 and Devetag, 2005).

- The possibility of observation of action choices, especially if paired with previous expressions of intent, is efficiency-enhancing (Duffy and Feltovich, 2002, 2006).

- Slowly growing groups that have managed to establish efficient precedents, is efficiency enhancing (Weber, 2005).

- Costly pre-play communication is efficiency-enhancing (e.g. VHBB, 1993).

- Costless pre-play communication is efficiency-enhancing (e.g., CDFR, 1992; Van Huyck et al., 1992; Blume and Ortmann, 2007; Duffy and Feltovich, 2002, 2006; Bangun et al., 2006).

- More meaningful communication, and common knowledge of information, are efficiency-enhancing (Chaudhuri et al., 2005; see also Bangun et al., 2006.)

- Loss avoidance may be efficiency-enhancing if losses are certain for a chosen action (e.g., Cachon and Camerer, 1996; Rydval and Ortmann, 2005; Feltovich et al., 2005).

5. Conclusion

We have qualitatively reviewed the evidence on coordination failure in the laboratory While two initial sets of experiments (VHBB, 1990, 1991; CDFR, 1990, 1992) seemed to suggest that coordination failure is almost inevitable, the sum total of subsequent attempts to understand the robustness of these results suggests myriad ways to engineer coordination successes in the lab.

Much of what we know about the emergence of coordination successes (and failures) in the lab seems related to what we have called structural determinants. We know surprisingly little about the impact of cognitive and behavioral determinants in these games. Since coordination is ultimately about trust (and since the literature on trust seems to mainly be concerned with behavioral and cognitive issues) the current state of affairs of research on coordination failures seems odd. Even elementary behavioral determinants such as the effects of risk attitudes have hardly been studied (see Heinemann et al., 2004a, for an important exception) although their potential impact has been indirectly acknowledged by some researchers analyzing stag-hunt games (e.g., the laudable but problematic early attempts by CDFR 1990, 1992 to control for risk preferences through the Roth – Malouf procedure in the stag hunt game)

Surprisingly, the impact of subject pools remains a blind spot (see Dufwenberg and Gneezy, 2005, and Cooper, 2006, for isolated exceptions).

Moreover, cognitive determinants (e.g., how subjects interpret and represent the payoff matrix) need to be investigated. There is tantalizing evidence (e.g., Cowan, 2001;

Devetag and Warglien, 2007; Wilcox, 1993) that the complexity of the matrix, and for that matter the task itself, systematically affects people's choice of strategies. We conjecture that, for example, the difference in results between Bangun et al. (2006) and Chaudhuri et al. (2005) is likely to reflect the complexities of the tasks involved. Somewhat surprisingly, there are no studies that use easily available tools such as MouseLab that have been used successfully in other contexts (e.g., Johnson et al., 2002) to study information acquisition and choice patterns. Also in need of investigation is the impact of precedent formation and transfer and the effects of both the quantity and quality of information.

Clearly, the question of how wide spread and pervasive coordination failure is in the lab, and in the wild, can hardly be answered conclusively by summarizing the extant experimental literature the way we have done. Ideally, one would start with an identification of a widely agreed-upon set of key determinants that would reflect the essential determinants in real-world situations of which order-statistic and stag-hunt games are claimed to be models of. Such a set of key determinants should span an agreed-upon parameter space which likewise ought to be informed by real-world settings (something which is a standard practice in macro economics but a practice essentially non-existent in micro-economics (NOTE9).)

Notwithstanding frequent appeals to real-world problems the coordination literature has not been much concerned with external validity, and definitely not with issues of calibration. Of course, not every experiment has to be calibrated. Much can be learned from experiments such as RVHB (2000) because they ask fundamental questions that are worth testing in the laboratory.

The evidence that we have reviewed above suggests myriad ways to engineer efficient outcomes in the lab. Most of these ways also seem to increase external validity (e.g., various forms of communication or repetition of slightly payoff perturbed games), at least for organizational contexts broadly construed. The potential of these efficiency-enhancing strategies to increase external validity for macro economic problems (and hence our ability to engineer efficient outcomes in the wild) is more difficult to assess.(NOTE10)

We conclude that, while coordination failures are common in the lab, they are by no means ubiquitous. More fundamentally, we are still far way away from an understanding of how common coordination failures are in the wild. We consider an answer to that question a challenge with high payoffs.

Acknowledgement

We thank John Duffy, Nick Feltovich, Jakub Steiner, two anonymous referees, and participants at the Coordination Success Conference in Honor of Raymond C. Battalio, College Station, March 31 – April 2, 2006, for their constructive comments. Special thanks to John Kagel for excellent comments on substance and style. The usual caveat applies. Significant parts of this manuscript were written while the authors visited the

Doctoral School of Economics and Management, CIFREM, of the University of Trento, Italy. We thank CIFREM for its hospitality.

Notes

1. A meta-study that quantitatively evaluates the impact of various factors on coordination (failure) is not (yet) possible since few authors have followed the advice of Davis and Holt (1993, p. 520) not to change too many things at once.
2. Space constraints forced significant selection on us.
3. This widely used terminology reflects the frequent interpretation of order-statistic games as problems of joint production. This terminology poorly reflects interpretations of these games as models of macro-economic activity. As a referee put it: “Keynes and macroeconomists think coordination failures arise from failures of price adjustment mechanisms (e.g. lots of people are unemployed/homes unsold because wages/home prices are sticky downward) - mechanisms that are not really present in any of these effort-based, game-theoretic interpretations.”
4. This confound was to some extent addressed in several later studies (e.g., Cachon and Camerer, 1996; Van Huyck et al., 2001; Goeree and Holt, 2005).
5. A risk-dominant equilibrium has a greater Nash product of deviation losses relative to the other equilibrium (e.g., Harsanyi and Selten, 1988). For example, the product of the deviation losses attached to the Pareto-inferior pure-strategy equilibrium is $(800-0)(800-0) = 640{,}000 > (1{,}000-800)(1{,}000-800) = 40{,}000$.

6. VHBB (1990, 1991) themselves conducted a number of important robustness tests. Among their key insights are the importance of the number of participants, the matching protocol, the feedback conditions, and the deviation costs.
7. In Blume and Ortmann (2007) pre-play messages take the form of "I intend to play action … ". Chaudhuri et al. (2005) allow for open-ended communication that they analyze for content, finding that subjects do not always focus on efficiency enhancing communication.
8. All statements below are ceteris paribus.
9. A laudable recent exception in the gift exchange literature is List (2006) who tackles the fundamental issue of the external validity of the laboratory evidence and the issue of calibration.
10. For starters there is the important discussion of what constitutes a large number in laboratory settings (e.g., Selten, 1973; Huck et al. 2004.)

REFERENCES

Aumann, R. (1990). "Nash Equilibria are Not Self-enforcing." Pp. 123 – 150 in J.J. Gabszewicz, J.-F. Richard and L.A. Wolsey (eds.). Economic Decision Making: Games, Econometrics and Optimization. Amsterdam: North_Holland.

Bangun, L., Chaudhuri, A., Prak, P. and Zhou, C. (2006). "Common and Almost Common Knowledge of Credible Assignments in a Coordination Game." Economics Bulletin, 3, 1-10

Battalio, R.C., Samuelson, L., and Van Huyck, J. (2001). "Optimization Incentives and Coordination Failure in Laboratory Stag Hunt Games." Econometrica, 69, 749-764.

Berninghaus, S.K. and Ehrhart, K.-M. (1998). "Time Horizon and Equilibrium Selection in Tacit Coordination Games: Experimental Results." Journal of Economic Behavior and Organization, 37, 231-248.

Berninghaus, S.K. and Ehrhart, K.-M. (2001). "Coordination and Information: Recent Experimental Evidence." Economics Letters, 73, 345-351.

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

Bornstein, G., Gneezy, U. and Nagel, R. (2002). "The Effect of Intergroup Comnpetition on Group Coordination: An Experimental Study," Games and Economic Behavior, 41, 1-25.

Brandts, J. and Cooper, D.J. (2004). "A Change Would Do You Good … An Experimental Study on How to Overcome Coordination Failure in Organizations." American Economic Review, 96, 669-693.

Brandts, J. and Cooper, D.J. (2006). "Observability and Overcoming Coordination Failure in Organizations." Experimental Economics, 9, 4, 407-423.

Burton, A., and Sefton, M. (2004). "Risk, Pre-play Communication and Equilibrium." Games and Economic Behavior, 46, 23-40.

Cachon, G.P. and Camerer, C.F. (1996). "Loss-avoidance and Forward Induction in Experimental Coordination Games." *Quarterly Journal of Economics,* 111, 165-194.

Camerer, C. (2003). Behavioral Game Theory. Experiments in Strategic Interaction. Princeton: Princeton University Press.

Charness, G. (2000). "Self-serving Cheap Talk: A Test of Aumann's Conjecture." Games and Economic Behavior, 33, 177-194.

Chaudhuri, A., Schotter, A. and Sopher, B. (2005). "Talking Ourselves to Efficiency: Coordination in Inter-Generational Minimum Effort Games with Private, Almost Common and Common Knowledge of Advice". Mimeo.

Clark, K., Kay. S. and Sefton, M. (2001). "When are Nash Equilibria Self-Enforcing? An Experimental Analysis." *International Journal of Game Theory*, 29, 435-456.

Clark, K. and Sefton, M. (2001). "Repetition and Signalling: Experimental Evidence from Games with Efficient Equilibria." Economics Letters, 70, 357 - 362.

Cooper, D.J. (2006). "Are Experienced Managers Experts at Overcoming Coordination Failure." Advances in Economic Analysis & Policy, 6.2, Article 6. [50 pages]

Cooper, R., De Jong, D., Forsythe, R., and Ross, T. (1990). "Selection Criteria in Coordination Games: Some Experimental Results." American Economic Review, 80, 218-233.

Cooper, R., De Jong, D., Forsythe, R., and Ross, T. (1992). "Communication in Coordination Games." Quarterly Journal of Economics, 107, 739-771.

Cowan, N. (2001). "The Magical Number 4 in Short-term Memory: A Reconsideration of Mental Storage Capacity." Behavioral and Brain Sciences, 24, 87-185.

Davis, D. & Holt, C.A. (1993). Experimental Economics. Princeton: Princeton University Press.

Devetag, G. (2005). "Precedent Transfer in Coordination Games: An Experiment." Economics Letters, 89, 227-232.

Devetag, G. and Warglien, M. (2007), "Playing The Wrong Game: An Experimental Analysis of Relational Complexity and Strategic Misrepresentation." Games and Economic Behavior, (forthcoming).

Dufwenberg, M. and Gneezy, U. (2005), "Gender and Coordination." Pp. 253-262 in A. Rapoport and R. Zwick (eds). Experimental Business Research, vol. 3. Boston: Kluwer.

Duffy, J. and Feltovich, N. (2002). "Do Actions Speak Louder Than Words? Observation vs. Cheap Talk as Coordination Devices." Games and Economic Behavior, 39, 1-27.

Duffy, J. and Feltovich, N. (2006). "Words, Deeds and Lies: Strategic Behavior in Games with Multiple Signals." Review of Economic Studies, 73, 3, 669-688.

Feltovich, N., Iwasaki, A. and Oda, S.H. (2005). "Payoff Levels, Equilibrium Selection, and Learning: An Experimental Study of the Stag Hunt." Mimeo.

Goeree, J.K. and Holt, C.A. (2005). “An Experimental Study of Costly Coordination.” Games and Economic Behavior, 51, 349–364.

Harsanyi, J. and Selten, R. (1988). A General Theory of Equilibrium Selection in Games, Cambridge, MA: The MIT Press.

Heinemann, F., Nagel, R. and Ockenfels, P. (2004a). “Measuring Strategic Uncertainty in Coordination Games.” Mimeo.

Huck, S., Norman, H., and Oechsler, J. (2004). “Two are few and four are many: number effects in experimental laboratories.” Journal of Economic Behavior and Organization, 53, 435–446.

Johnson, E.J., Camerer, C., Sen, S. and Rymon, T. (2002). “Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining.” Journal of Economic Theory, 104, 16–47.

Knez, M. and Camerer, C. (1994). “Creating Expectational Assets in the Laboratory: Coordination in 'Weakest-Link' Games.” Strategic Management Journal, 15, 101-119.

Knez, M. and Camerer, C. (2000). “Increasing Cooperation in Prisoner's Dilemmas by Establishing a Precedent of Efficiency in Coordination Games.” Organizational Behavior and Human Decision Processes, 82, 194-216.

Kreps, D.M. (1990). Game Theory and Economic Modelling (Clarendon Lectures). Oxford: Oxford University Press.

Lewis, D. (1969). Convention: A Philosophical Study. Cambridge: Harvard University Press.

List, J.A. (2004). “Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace.” Econometrica, 72, 615-625.

List, J. A. (2006). “The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions.” Journal of Political Economy, 114, 1, 1-37.

Plott, C.R. and Zeiler, K. (2005). “The Willingness to Pay-Willingness to Accept Gap, the ‘Endowment Effect’, Subject Misconceptions and Experimental Procedures for Eliciting Valuations.” American Economic Review, 95, 530-545.

Rankin, F., Van Huyck, J.B., and Battalio, R.C. (2000). “Strategic Similarity and Emergence of Conventions: Evidence from Payoff Perturbed Stag Hunt Games.” Games and Economic Behavior, 32, 315-337.

Rydval, O. and Ortmann, A. (2005). "Loss Avoidance as Selection Principle: Evidence from Simple Stag-hunt Games." Economics Letters, 88, 101-107.

Schmidt, D., Shupp, R., Walker, J.M., and Ostrom, E. (2003). "Playing Safe in Coordination Games: The Role of Risk Dominance, Payoff Dominance, Social History, and Reputation." Games and Economic Behavior, 42, 281-299.

Selten, R. (1973). "A simple model of imperfect competition, where four are few and six are many." International Journal of Game Theory, 2, 141-201.

Van Huyck, J. B., Battalio, R.C., and Beil, R.O. (1990). "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure." The American Economic Review, 80, 234-248.

Van Huyck, J. B., Battalio, R.C., and Beil, R.O. (1991). "Strategic Uncertainty, Equilibrium Selection, and Coordination Failure in Average Opinion Games." The Quarterly Journal of Economics, 106, 885-911.

Van Huyck, J. B., Battalio, R.C., and Beil, R.O. (1993). "Asset Markets as an Equilibrium Selection Mechanism: Coordination Failure, Game Form Auctions, and Tacit Communication." Games and Economic Behavior, 5, 485-504.

Van Huyck, J.B., Battalio, R.C., and Rankin, F.W. (2001). "Evidence on Learning in Coordination Games." Texas A&M University laser script.

Van Huyck, J.B., Cook, J.P. and Battalio, R.C. (1997). "Adaptive Behavior and Coordination Failure." Journal of Economic Behavior and Organization, 32, 483-503.

Van Huyck, J.B., Gillette, A. and Battalio, R.C. (1992). "Credible Assignments in Coordination Games." Games and Economic Behavior, 4, 606-626.

Weber, R. (2005). "Managing Growth to Achieve Efficient Coordination in Large Groups." American Economic Review, forthcoming.

Weber, R., Camerer, C., Rottenstreich, Y. and Knez, M. (2001). "The Illusion of Leadership: Misattribution of Cause in Coordination Games." Organization Science, 12, 582-598.

Wilcox, N.T. (1993). "Lottery Choice: Incentives, Complexity and Decision Time." Economic Journal, 103, 1397-1417.