Behavioral Forces Driving Information Unraveling

Volker Benndorf (Goethe-Universität Frankfurt)
Dorothea Kübler (WZB Berlin, TU Berlin and CESifo)
Hans-Theo Normann (Universität Düsseldorf)

Discussion Paper No. 354

December 22, 2022

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Behavioral Forces Driving Information Unraveling*

Volker Benndorf† Dorothea Kübler‡ Hans-Theo Normann§

May 2022

Abstract

Information unraveling is an elegant theoretical argument suggesting that private information may be fully and voluntarily surrendered. The experimental literature has, however, failed to provide evidence of complete unraveling and has suggested senders' limited depth of reasoning as one behavioral explanation. In our novel design, decision-making is essentially sequential, which removes the requirements on subjects' reasoning and should enable subjects to play the standard Nash equilibrium with full revelation. However, our design also facilitates coordination on equilibria with partial unraveling which exist with other-regarding preferences. Our data confirm that the new design is successful in that it avoids miscoordination entirely. Roughly half of the groups fully unravel whereas other groups exhibit monotonic outcomes with partial unraveling. Altogether, we find more information unraveling with the new design, but there is clear evidence that other-regarding preferences do play a role in impeding unraveling.

JEL Classification numbers: C72, C90, C91

Keywords: data protection, inequality aversion, information revelation, level-k reasoning.

*We are grateful to seminar and conference audiences at BCCP Berlin, Göttingen, Grenoble, Liverpool, Paris (CREST) and Royal-Holloway College for helpful discussions and comments. Jennifer Rontganger provided excellent copy-editing service. Financial support from the DFG through CRC TRR 190 (Kübler) is gratefully acknowledged.

†Corresponding author, Goethe-Universität Frankfurt, Theodor-W.-Adorno-Platz 4, HPF 84, 60629 Frankfurt, Tel: +49 69 79834806, email: *benndorf@econ.uni-frankfurt.de.* Declarations of interest: none

‡WZB Berlin and Technical University Berlin, Tel: +49 30 25491 440; email: *kuebler@wzb.eu.* Declarations of interest: none

§Duesseldorf Institute for Competition Economics (DICE), Tel: +49 211 8115297, email: *normann@dice.hhu.de.* Declarations of interest: none

1 Introduction

George Akerlof commented on the, at that time, new car insurance policies with voluntary GPS tracking ("Black Box") as follows:[1]

> *"It will be interesting to see what will happen. ... When the black box becomes more widespread, it will be mainly those drivers who drive carefully anyway who will buy one. They hope to be able to lower their insurance premiums. The others will continue to drive without a box. Insurance for cars without a black box will become more and more expensive because the insurance companies know that they tend to be the worse risks. People who don't want to buy a black box ... may eventually no longer be able to get car insurance at all."*

The quote succinctly summarizes the logic behind the information unraveling process as put forward in the theoretical literature (Viscusi, 1978; Grossman and Hart, 1980; Grossman, 1981; Milgrom, 1981; Milgrom and Roberts, 1986). Privately informed players may completely and voluntarily disclose verifiable information. In a (hypothetical) dynamic thought process, initially only some senders, namely those with the most favorable information, have an incentive to reveal their private information. Given these players reveal, others will also find it profitable to reveal, which makes revelation profitable for even more players, and so on. In the end, only the players with the least favorable information continue to conceal their private information. But, since the concealing players are identified by the fact that they do not disclose, information unraveling is complete. Rational senders anticipate the result of this thought process and immediately reveal all their information in the one-shot game.

The quote, however, also highlights that decisions in the field may not immediately unravel like the theory stipulates. In the field, players do not decide simultaneously in a one-shot game; the iterations of the hypothetical dynamic thought process are real sequential moves: Some senders reveal their information first whereas others follow only after they have observed changes in their payoffs. Players decide sequentially, or at least they have various opportunities where they can reveal. Even if they reveal eventually, it may take them several iterations to decide.

The distinction between immediate complete revelation in theory and a slow and somewhat impeded decision-making process in practice may be important when it comes to the interpretation of the evidence. A growing number of experimental papers take models with information unraveling to the lab (Benndorf, Kübler, and Normann, 2015; Benndorf, 2018; Hagenbach and Perez-Richet, 2018; Penczynski and Zhang, 2018; Jin, Luca, and Martin, 2021). Decision-making in these experiments corresponds to the simultaneous-move setup of the theory.[2] A common theme connecting these papers is that, by and large, information unraveling is incomplete (see Section 2 below). Senders do not fully disclose information and it is not only the players with the least favorable types that decide to conceal. Does the simultaneous-move setup overburden players and preclude complete unraveling?

In Benndorf, Kübler, and Normann (2015), we indeed argue that limited depth of reasoning can explain incomplete revelation by senders. The level-k model (Nagel, 1995) nicely fits with the behavioral patterns we observe in unraveling experiments. The revelation decisions of players with favorable information require little or no high-level reasoning about others' decisions. As players with not-so-favorable information (who should, in theory, nevertheless reveal) need to anticipate the behavior of others at higher levels of reasoning, they are

[1] Akerlof quoted and paraphrased in "Revolution in der KfZ Versicherung," Frankfurter Allgemeine Zeitung, 13/01/2014, `https://www.faz.net/-ht4-7l81d`, last retrieved 07/03/2022.

[2] In Jin, Luca, and Martin (2021), Penczynski and Zhang (2018), and Hagenbach and Perez-Richet (2018), decisions making is with simultaneous moves and with random rematching after each round. Benndorf, Kübler, and Normann (2015) and Benndorf (2018) have simultaneous moves with fixed matching. Neither design corresponds to the quasi-sequential decision-making we introduce in this paper.

more prone to making "wrong" decisions and conceal their information. After all, revelation is profitable for them only conditional on players with favorable information revealing. We find that the more steps of reasoning a player has to think through, the less likely she is to reveal. It is therefore not surprising that we observe incomplete revelation in the lab.

The main innovation in this paper is an experimental design that does not require great sophistication for information unraveling to occur. While previous experiments involve simultaneous decision-making, in our novel design, decision-making is "quasi sequential." Participants have five minutes to decide, and during this period they see the current profile of decisions on the computer screens. Subjects do not need to anticipate others' decisions any more. They can wait and observe what other players (perhaps those whose decisions require a low level-k) do before deciding themselves.[3] The decision to reveal information simply boils down to a comparison of the two payoffs resulting from reveal/conceal decisions. Others' decisions do not have to be anticipated, and so revelation decisions do not require higher-level reasoning. In terms of the level-k model, senders only need a level of $k = 1$ to reveal. In fact, decision-making in quasi-sequential designs bear a closer resemblance to the slow step-by-step process the quote characterizes rather than the simultaneous-move one-shot game the theory analyzes. Having said that, the primary purpose of the novel design is not to improve realism, but to remove the strategic uncertainty.[4]

It may appear that, if quasi-sequential decision-making is successful at removing level-k requirements, senders will always fully reveal their information, but this is not necessarily the case. Our quasi-sequential design not only purges any requirements on depth of reasoning, it may at the same time facilitate the coordination on equilibria. With standard preferences the unique Nash outcome involves complete unraveling, but with other-regarding preferences (Fehr and Schmidt, 1999, henceforth "FS"), several equilibria with incomplete information disclosure exist. If senders with favorable information anticipate the unraveling process, they may refrain from disclosing their information because doing so would impose a negative externality on other players. In fact, a calibrated model with rational FS players shows that any extent of partial unraveling can be an equilibrium. With simultaneous decision-making and due to multiple equilibria, players may find it difficult to coordinate on any of these FS equilibria. In our setup, they have ample time to coordinate, seeing others' decisions and the impact of their own behavior on the screen. Hence, we might see more (rather than fewer) outcomes with incomplete information disclosure because of the improved coordination of the new design.

In two further treatments we subject inequality aversion to a stress test. We expect both treatments to have an effect in the same direction, but we decided to implement them both because of their potential relevance in the field.[5] In one treatment we induce entitlement. Some players may feel they deserve higher payoffs because they won a pre-game contest (see, for example, Fahr and Irlenbusch, 2000; Konow, 2000; Oxoby and Spraggon, 2008; Fischer and Normann, 2019). Such entitlement is likely to be a relevant factor in the field and entitlement suggests information revelation, and it weakens inequality aversion. In the insurance examples, people may feel entitled to a better contract because they do more sports, drive more carefully etc. We hypothesize that such entitlement reduces fairness concerns which should trigger further unraveling. In another treatment, our participants are unaware of the externality unraveling imposes on others.[6] Other-regarding motives can only kick in if players understand the impact information revelation has on other players. This may not be given

[3] Note that we do ensure that subjects in the experiment cannot surprise other subjects by deciding in the final seconds of the five-minute period. See Section 4 for details.

[4] Decisions that are actually sequential require an extensive-form game with a fixed order of moves. In our quasi-sequential design, players can coordinate without such an exogenously given sequence of moves and payoffs materialize only once, at the end of the five-minute period.

[5] Both treatments can be related to the above quote. Those who "drive more carefully" will see little reason to subsidize reckless drivers. And the externality, is observable only later, after the first movers have decided ("[i]nsurance ... will get more expensive ... for higher risks").

[6] To be precise, we do not explicitly inform subjects about the externality, but they can still deduce the impact of the decisions from the information they have on the computer screen.

in the field where the externality is presumably less transparent, also due to the framing of decisions. In the insurance examples, policy holders merely get a “bonus” for good behavior (exercise regularly, drive carefully) while premiums for others policy holders stay constant in the short run, so the negative externality is not obvious.[7] When participants fail to recognize the externality, as may be the case in this treatment, this leads to further unraveling.

To summarize, we analyze four experimental treatments which focus on the unraveling decisions of senders of private information. We implement the receivers via automatic and computerized decisions.[8] Except for a baseline variant with simultaneous decisions and comparable to previous experiments, our new treatments use quasi-sequential decision-making. Given our design is successful at removing requirements regarding players' depth of reasoning, any remaining information concealment is likely due to preferences for payoff equality. Two further treatments should then reduce the impact of other-regarding preferences.

Our research questions are: Does limited depth of reasoning impede unraveling? Does quasi-sequential decision-making successfully remove requirements on the depth of reasoning, such that we observe improved coordination between players? Furthermore, does quasi-sequential decision-making lead to more, possibly full information revelation? Is the timing of decisions in the experiment consistent with the hypothetical thought process the theory stipulates? Last but not least, how much explanatory power does inequality aversion have?

The results show that our novel design with quasi-sequential decision-making was successful in that we observe 100% equilibrium outcomes—compared to slightly more than half of the groups in the standard design. That is, players coordinate much better on monotonic outcomes that correspond to either the standard Nash or an FS equilibrium. We see full unraveling (we believe, for the first time in experiments) in roughly half of the groups. The remaining groups show different degrees of incomplete revelation, consistent with inequality concerns. The timing of decisions suggests that revelation behavior is indeed consistent with the hypothetical thought process Akerlof described in his quote. Average revelation rates are higher with quasi-sequential decision-making compared to a baseline variant with simultaneous decision-making. Inequality aversion does play a role: The two further treatments designed to reduce inequality concerns indeed lead to more revelation but only insignificantly so.

2 Related literature

The literature on information revelation can be roughly divided into three areas: signaling,[9] cheap talk,[10] and—our focus—the disclosure of verifiable information. Experiments in this area build on an established body of theoretical works (Viscusi, 1978; Grossman and Hart, 1980; Grossman, 1981; Milgrom, 1981; Milgrom and Roberts, 1986).

The first experimental paper on verifiable information revelation is presumably Forsythe, Isaac, and Palfrey, 1989. They study “blind bidding” in the motion picture industry. In their experiment, sellers have private

[7] Gächter et al. (2009) show that even experimental economists are prone to such a “bonus” versus “penalty” framing.

[8] Various behavioral factors may impede the information-unraveling process and bounded rationality may occur on both the sender and the receiver side. Already, Milgrom and Roberts (1986) conjectured that buyers of a good with, say, uncertain product quality may not br sophisticated enough to understand that “no news is bad news” and that this may induce—potentially more sophisticated—firms (senders) to reveal only positive information (see the papers discussed in Section 2).

[9] In signaling games, senders may signal private information with a distorted action (Spence, 1973), and equilibria may convey information to the receiver. Experiments in this area include Miller and Plott (1985), Cadsby, Frank, and Maksimovic (1990), Cadsby, Frank, and Maksimovic (1998), Brandts and Holt (1992), Brandts and Holt (1993), Cooper, Garvin, and Kagel (1997b), Cooper, Garvin, and Kagel (1997a), Potters and Van Winden (1996), Cooper and Kagel (2003), and Kübler, Müller, and Normann (2008)

[10] In papers on cheap talk and persuasion, the informed player can costlessly communicate anything (Crawford and Sobel, 1982). Laboratory experiments in this second realm include Isaac and Walker (1988), Bochet, Page, and Putterman (2006), Blume and Ortmann (2007), Brandts and Cooper (2007), Andersson and Wengström (2007), Bochet and Putterman (2009), Fonseca and Normann (2012), Oprea, Charness, and Friedman (2014), and Moellers, Normann, and Snyder (2017). These experiments are surveyed by Crawford (1998) and Balliet (2010).

information about a good and decide whether to reveal this information to the buyers. Revelation costs are zero. Participants learn to largely reveal private information over 16 to 22 rounds of play, but revelation was still incomplete (see their Table 2 with "contradictory observations to the unraveling process").[11]

In our previous work (Benndorf, Kübler, and Normann, 2015), revelation disclosure is costly. Six informed players (framed as workers) have to decide whether to reveal their productivity (framed as the purchase of a health certificate) to uninformed parties (employers) where the latter were not human subjects but automated computer moves. We find that revelation rates are too low compared to the prediction. In our main variant, we observe only slightly more than 50% revelation, compared to the prediction of 83.3%. While a non-loaded frame leads to 10 percentage points increase in revelation, altogether the data do not seem to converge to full revelation in any treatment.

Benndorf (2018) uses the same parameters as Benndorf, Kübler, and Normann (2015), but includes human players as the receivers (employers). The study supports the earlier results in that revelation rates are too low compared to the predictions. As for the receivers, Benndorf (2018) points out that employers do not make positive profits when employing workers of unknown productivity whereas they do make positive profits when employing workers who revealed their productivity. The latter effect is inconsistent with the equilibrium predictions and may further mitigate the information disclosure of the senders. Another finding is that reducing the costs of revelation to a negligible degree does not necessarily imply a higher degree of equilibrium-consistency. This may be interpreted as further evidence that a limited depth of reasoning mitigates the unraveling of private information.

In the experiment of Jin, Luca, and Martin (2021), both parties are represented by human participants. Their research question is whether "no news is bad news," that is, whether receivers are sufficiently pessimistic about senders who choose not to disclose. It turns out they are not. This insufficient pessimism reduces senders' incentives to reveal: senders disclose favorable information, but withhold less favorable information. Feedback on the interactions helps players learn and reach equilibrium.

Penczynski and Zhang (2018) study information unraveling in a competitive setting. Even when receivers are not sufficiently sophisticated, competition between informed senders will lead to provision information unraveling (Milgrom and Roberts, 1986). Penczynski and Zhang (2018) compare the competitive setting to monopoly. They also find that buyers are not skeptical enough, and that this is particularly the case in the competition treatment. Another competitive setting is explored in Ackfeld and Güth (2019) who also extend the literature on information disclosure to the case where the information is actually privacy sensitive. For a duopoly with behavior-based pricing, Heiny, Li, and Tolksdorf (2020) find that consumers reveal their private data in about two thirds of cases, confirming that information revelation is incomplete. Güth et al. (2019) analyze in theory and experiment the case of welfare-enhancing information revelation in an acquiring-a-company game. An experiment by Hagenbach and Perez-Richet (2018) investigates non-monotonic incentives whereas the aforementioned literatures studies situations where senders have monotonic incentives: they prefer to be perceived as having higher productivity, better quality, etc. Hagenbach and Perez-Richet (2018) also include games with non-monotonic and cyclical incentives. Their data are consistent with a non-equilibrium model based on the iterated elimination of obviously dominated strategies.

There are also field studies that suggest information unraveling is incomplete. See Luca and Smith (2013) for a study where business colleges only publicized rankings in which they did well. According to Bederson et al. (2018) restaurants often did not reveal their hygiene standards unless required to. In Mathios (2000), many grocery store food items did not include nutritional information until it was mandated. Frondel, Gerster, and Vance (2020) find that house owners reduce offer prices only when the disclosure of energy efficiency information

[11] Dickhaut, McCabe, and Mukherji (1995) show that, the more aligned the interests of the sender and receiver are, the more informative the sender's message turns out to be.

is mandatory, and that the effect is stronger for owners who did not disclose when it was not mandatory.

3 The Game

3.1 The game

There are n players who we refer to as *workers*. Workers are heterogeneous with respect to their productivity θ_i. Productivities are ordered such that $\theta_1 \geq \theta_2 \geq \cdots \geq \theta_n$. All workers have two actions, they can either *reveal* their productivity to a fictitious employer, or they can *conceal* (not reveal) it. The decision of worker i is denoted by I_i where $I_i = 1$ indicates revelation whereas $I_i = 0$ denotes concealment. Revealing involves a cost, c, whereas concealing is free.

Workers are paid according to the following payoff function

$$\pi_i = I_i \left[\theta_i - c\right] + (1 - I_i) \left[\frac{\theta_i + \sum_{j \neq i}(1 - I_j)\theta_j}{1 + \sum_{j \neq i}(1 - I_j)}\right] \tag{1}$$

In words, workers who reveal receive their productivity as a wage payment minus the cost of revelation, c. Workers who conceal receive the average productivity of all concealing workers as a wage payment but do not pay c.

3.2 Nash equilibrium

In Benndorf, Kübler, and Normann (2015), we show that $I_1 \geq I_2 \geq \cdots \geq I_n = 0$ in any Nash equilibrium. Which players reveal or conceal in equilibrium depends on the set of possible productivities $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ and the cost of revelation c. In the experiment, we have $n = 6$, $c = 100$, and $\Theta = \{607, 582, 551, 510, 448, 200\}$ and get a unique Nash equilibrium with standard preferences where $I_1 = \cdots = I_5 = 1$ and $I_6 = 0$, that is workers fully reveal their information.

3.3 Monotonic outcomes, Fehr-Schmidt equilibria, and level-k rationality

3.3.1 Monotonic outcomes

Even though the Nash equilibrium of the game with standard preferences is unique, there are further strategy profiles that are relevant in the experiment. For the parameterization we use, there are $2^6 = 64$ pure strategy combinations six of which are monotonic in that $I_1 \geq I_2 \geq \cdots \geq I_6 = 0$. We label the monotonic outcomes M1 to M6 where the labels refer to the number of players who conceal. M1 with full revelation is the Nash equilibrium with standard preferences. Table 1 provides an overview.

Worker	M1	M2	M3	M4	M5	M6
1	1	1	1	1	1	0
2	1	1	1	1	0	0
3	1	1	1	0	0	0
4	1	1	0	0	0	0
5	1	0	0	0	0	0
6	0	0	0	0	0	0

Table 1: Monotonic outcomes of the revelation game. The entries "0" and "1" for worker i indicate $I_i = 0$ and $I_i = 1$, respectively. The label "Mj" indicates that workers 1 to j conceal in that outcome. M1 is the standard Nash equilibrium.

Both the FS and the level-k model suggest that the monotonic outcomes are more likely to be reached than non-monotonic ones, even though there are far more non-monotonic outcomes. We now discuss both models in turn.

3.3.2 Fehr-Schmidt equilibria

The FS model of inequality aversion is based on the following utility function

$$U_i(x_1, ..., x_n) = x_i - \frac{\alpha_i}{n-1} \sum_{j \neq i} \max\{x_j - x_i, 0\} - \frac{\beta_i}{n-1} \sum_{j \neq i} \max\{x_i - x_j, 0\} \tag{2}$$

where x_i and x_j, $i \neq j$, denote the payoffs to i and j. FS assume $\beta_i \leq \alpha_i$ and $0 \leq \beta_i < 1$.

The intuition for why inequality aversion may lead individual players to conceal is straightforward. Consider the profile where workers 2 to 5 conceal. If worker 1 also conceals, everybody earns 483 and there is zero inequality. If worker 1 reveals, she earns 507 while the remaining workers all get 458.2. It follows that worker 1 prefers to conceal if $483 \geq 507 - 48.8\beta_1$ or $\beta_1 \geq 24/48.8 \approx 0.492$. Given worker 1 is sufficiently averse towards advantageous inequality, she will not reveal.

For the parameters used in the experiment, the outcomes presented in Table 1 may all be FS equilibria in pure strategies. More generally, we prove in Appendix B the following proposition:

Proposition. The monotonic strategy profiles listed in Table 1 are FS equilibria for some inequality-aversion parameters. There are no other FS equilibria in pure strategies.

Whether some monotonic outcome is actually an FS equilibrium depends on the α_i-β_i parameters. Table 4 in Appendix B lists the specific conditions for each outcome. The table also shows how likely these conditions are met based on the joint distribution of α_i-β_i parameters observed in Blanco, Engelmann, and Normann (2011). The intuition for why so many FS equilibria exist (as long as they are monotonic) is similar to above for worker 1: Given $I_j = 1$ for all $j < i$ and $I_j = 0$ for all $j > i$, the FS parameters of worker i determine whether i will reveal or conceal.

The core part of the proof of the second part of the proposition (see Lemma 2 in Appendix B) boils down to generally showing that, given concealing is the best reply with standard preferences, concealing is also the best reply when players are inequality averse. This suggests that, somewhat loosely speaking, inequality aversion cannot lead to more revelation. As we show in Appendix B, this also implies that no non-monotonic outcomes can be FS equilibria.

3.3.3 Level-k model

We now analyze the revelation game with the level-k model of limited depth of reasoning. We assume that $k = 0$ players randomize uniformly between their two actions. A $k = 1$ player believes all other players are $k = 0$. This changes incentives if she is worker 1: Straightforward calculations show that worker 1 best responds by revealing. If the $k = 1$ player is any other worker $j > 1$, she conceals. Next, a $k = 2$ player assumes all other workers are level $k = 1$ and hence that worker 1 will reveal. Conditional on worker 1 revealing (due to $k = 1$), the $k = 2$ player will reveal as worker 2, but she will conceal as worker $j > 2$. If the $k = 2$ player is worker 1, she will likewise reveal. Generally, worker j, $j \leq 5$, will reveal if and only if $k \geq j$ for worker j. Worker 6 has a strictly dominant action and will conceal for any $k \geq 1$.

In a simultaneous-move one-shot game, limited depth of reasoning can give rise to both monotonic and non-monotonic outcomes. Consider the profile M3 (workers 1 to 3 reveal and workers 4 to 6 conceal). This profile will occur if the following constraints on k are met: Worker 1 $k \geq 1$, worker 2 $k \geq 2$, worker 3 $k \geq 3$, worker 4 $k < 4$, worker 5 $k < 5$, and worker 6 $k \geq 1$. Contrast this with the similar but non-monotonic profile where workers 1 and 2 reveal, worker 3 conceals, worker 4 reveals, and workers 5 and 6 conceal. This profile will occur if and only if we have worker 3 $k < 3$ and worker 4 $k \geq 4$ (and under the same conditions as in the monotonic example for the other workers). Generally, a non-monotonic outcome will arise when some worker j, $j \geq 2$, has level $k \geq j$ whereas some worker $i < j$ has level $k < i$.

The k-levels of laboratory participants can be quite heterogeneous (Arad and Rubinstein, 2012; Benndorf, Kübler, and Normann, 2017), but this does not mean that monotonic outcomes will be rare. Randomly drawing individual k-levels from the distribution of k-levels reported in Arad and Rubinstein (2012) and Benndorf, Kübler, and Normann (2017),[12] the above monotonic outcome will occur in 19.8% of the cases whereas the non-monotonic outcome will occur in 8.4%.

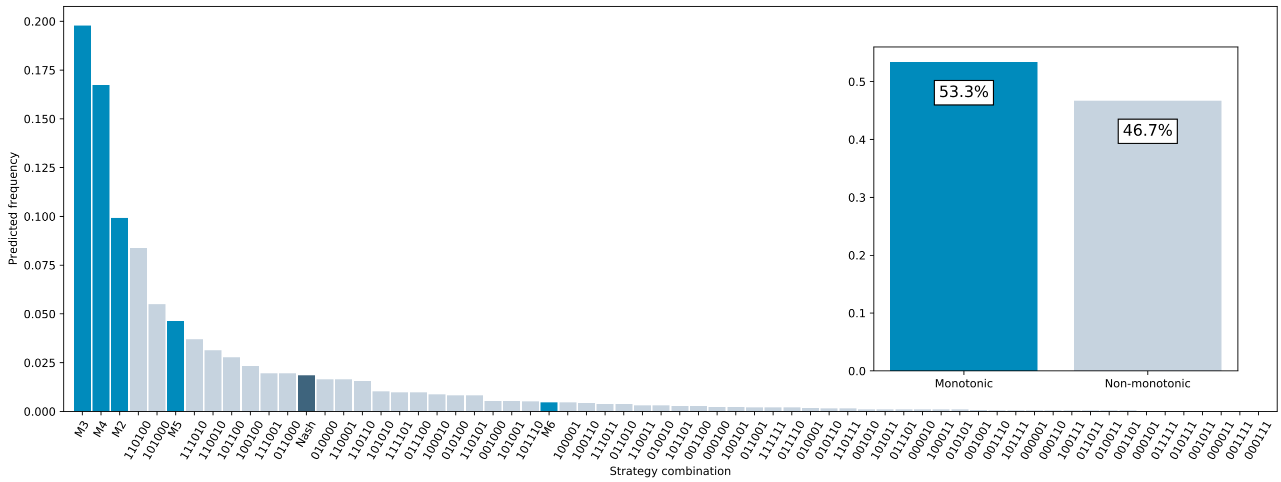

Figure 1: Frequencies of group-level outcomes as predicted by the level-k rationality model. The inset shows the aggregated probabilities of observing a monotonic or non-monotonic outcome.

Applying this logic to all 2^6 possible outcomes unveils that the level-k model actually favors monotonic outcomes, see Figure 1. The figure shows that, based on the data in Arad and Rubinstein (2012) and Benndorf, Kübler, and Normann (2017) we can expect about 53.3% monotonic outcomes and 46.7% non-monotonic outcomes. This calibration suggests that the standard Nash equilibrium with full revelation is unlikely to occur (less than 2%), but the all-conceal outcome is even less likely. The main takeaway from the figure is that, even without explicit coordination between workers, slightly more than half of the outcomes will be monotonic.

4 Experimental design and procedures

4.1 Basic design

In all treatments, the experiments consist of two parts, a real-effort task followed by the game described in the previous section. Both parts are played only once, so this is a one-shot experiment.

In the real-effort task, subjects had to "encode" so-called "words" (consisting of three random letters) into three-digit numbers. The mapping from letters to numbers is presented on the subject's computer screen and

[12] The cumulative distribution functions of k-levels Arad and Rubinstein's (2012) Basic and Benndorf, Kübler, and Normann's (2015) Market B (first period) are virtually identical. Our calibration is based on Table 2 (without conceal decisions) in Benndorf, Kübler, and Normann (2015).

changes every time a word has been encoded correctly.[13] Subjects had 10 minutes to encode as many words as they like. In one treatment, performance in the real-effort task was used to allocate productivities in the second part. In all other treatments, the first part did not play a role for the game part.

In the game part, subjects play the revelation game as described above in Section 3.1. The parameters we used are $n = 6$, $c = 100$, and $\Theta = \{607, 582, 551, 510, 448, 200\}$, as in Market B of Benndorf, Kübler, and Normann (2015).

In all treatments (except for the "Baseline" control variant), subjects decide quasi-sequentially in continuous time. The initial (default) decision is to conceal. Subjects then have five minutes to choose to reveal or switch back to conceal. During these five minutes, they may repeatedly change their decision. Only the decision at the end of that period is relevant for subjects' payoffs. We wanted to avoid surprise changes of decisions at the very end of the period. So, if anybody changed their decision in the final 10 seconds of the five-minute interval, the decision-making period was extended for another 10 seconds.

4.2 Treatments

We consider four treatments in a between-subjects design. The treatments are labeled *Baseline*, *Sequential*, *Entitlement*, and *Unaware* (or B, S, E, U). We vary the information participants have when they make their revelation decision and whether or not participants' productivities in the revelation game are tied to their performance in the real-effort task. See Table 2 for an overview. A reproduction of the decision screens is provided in Figure 5 in Appendix A.

In Baseline, subjects learn the payoff function from the instructions, and they are informed about their own productivity and the set of possible productivities when they make their revelation decision. They do not receive any new information during the decision-finding process. Subjects' productivities are determined by a random draw of the computer. The random draw is without replacement such that each group comprises exactly one worker of each type. decision-making is with simultaneous moves. In summary, the Baseline treatment is a standard game with strategic uncertainty.

In Sequential and in Entitlement, subjects have all the information they have in the Baseline treatment, but they additionally see the currently selected strategy profile of their group. The also see payoffs resulting from either of their actions implied by the current strategy profile. More precisely, subjects see a list of all six workers in their group (including themselves), the productivities of these workers, and the current revelation choice of each worker. Decision-making is quasi-sequential: Whenever a worker changes her decision, this information is immediately relayed to all other participants in the group. This procedure effectively removes all strategic uncertainty, such that erroneous beliefs or miscoordination cannot play a role. Since subjects are also informed about the resulting profits, decision errors based on incorrect calculation can also be ruled out.

The key difference between Sequential and Entitlement is the way the productivities in the revelation game are assigned to participants (Table 2). In Entitlement, the assignment is tied to subjects' performance in the real-effort task: the worker with the highest productivity will be the subject who encoded the most words in the real-effort task, the second highest productivity will be the subject who encoded the second-most words, and so on. In Sequential and Baseline, this role assignment is determined by a random computer draw. The idea behind the treatment manipulation in Entitlement is that tying the role assignment to real-effort performance may mitigate other-regarding motives such as inequality aversion or joint-surplus maximization. Participants may feel more inclined to maximize their own profits because they performed better, even though this imposes a negative externality on other players in their group.

[13]The task was designed to prevent subjects from getting better at doing the task over time. This is why the mapping from letters to numbers and the order of the letters in the table change after each word. See Benndorf, Rau, and Sölch (2019b) and Benndorf, Rau, and Sölch (2019a) for more details.

Table 2: Treatments

	Information available during revelation decision	Productivity	Decision-making
Baseline	Payoff function, own productivity and set of possible productivities	Random	Simultaneous
Sequential	Payoff function, own productivity, set of possible productivities, strategy profile in group, and resulting payoffs from both actions	Random	Quasi-sequential
Entitlement	Payoff function, own productivity, set of possible productivities, strategy profile in group, and resulting payoffs from both actions	Real-effort performance	Quasi-sequential
Unaware	Own productivity, set of possible productivities, and payoffs from both actions	Random	Quasi-sequential

The fourth treatment, called Unaware, is designed to reduce subjects' awareness of the negative externality they exert on others when they reveal. Different from Sequential, the instructions do not explain in detail the profit function when concealing. Subjects only learn that their concealment profits will depend on the behavior of the other workers in their group. Furthermore, the decision screen does not display any information about the other members of the group or their behavior. It does, however, display the subject's two potential profits given the current strategy profile of the group, so subjects can still learn the payoff function. So, subjects are still able to identify optimal behavior even though the exact payoff function is ex-ante unknown.

4.3 Procedures

The experiment was conducted at the DICELab at Düsseldorf university. Most participants were students, but there were also some university employees. We conducted a total of nine sessions, each with 24 to 36 participants (that is, four to six groups of six). A total of 276 subjects participated in this study. The sessions were conducted between February 2016 and June 2018 using the usual combination of z-Tree and ORSEE (Fischbacher, 2007; Greiner, 2015). Sessions lasted about 45 minutes and the average payment was 10.52 Euros which includes a show-up fee of 2 Euros.

5 Hypotheses

As discussed in the theory section, all equilibria of the game (with and without other-regarding preferences) are monotonic. Hence, non-monotonic outcomes are coordination failures which should not occur in the treatments where decision finding is quasi-sequential. In contrast, a calibrated level-k model suggests that the share of non-monotonic outcomes in Baseline might be around 46.7% (see Section 3.3.3 and Figure 1). We therefore hypothesize that we will observe more non-monotonic outcomes in Baseline compared to all other treatments.

Hypothesis 1. *There will be more non-monotonic outcomes in Baseline compared to all other treatments.*

A robust finding of the level-k literature is that relatively many subjects reason at k-levels of up to three whereas higher levels of reasoning are hardly reported (see Bosch-Domenech et al., 2002; Arad and Rubinstein, 2012; Benndorf, Kübler, and Normann, 2017). Since the Nash equilibrium of the game with standard preferences requires k-levels up to five, it follows that Nash outcomes with complete unraveling are unlikely to be reached in the simultaneous-move setup in Baseline (see Figure 1 above). Sequential, Entitlement, and Unaware do not require higher levels of reasoning such that the Nash equilibrium may actually be observed in these treatments.

Hypothesis 2. *There will be fewer outcomes with full unraveling in Baseline compared to all other treatments.*

As for the degree of information revelation in general, the total effect of removing strategic uncertainty in Sequential is ambiguous. On the one hand, there might be more unraveling in Sequential compared to Baseline as limited depth of reasoning no longer impedes revelation of low-productivity workers. On the other hand, there might also be less unraveling. The quasi-sequential decision-making enables subjects to coordinate on FS equilibria with incomplete unraveling (M2-M6). One of them (M6) is not only the second-most robust equilibrium,[14] but also implies zero revelation and it is the default choice at the beginning of the decision-making process. Altogether, this suggests that there might be more or less unraveling in Sequential compared to Baseline.

Hypothesis 3a. *There will be more revelation in Sequential compared to Baseline.*

Hypothesis 3b. *There will be less revelation in Sequential compared to Baseline.*

Compared to Sequential, both Entitlement and Unaware reduce aversion toward inequality and should thus lead to more revelation. Entitlement is identical to Sequential except that the allocation of productivities is not random but according to a contest which should weaken other-regarding motives.[15] In Unaware, subjects were not explicitly told about the negative externality revelation imposes. Inequality-averse subjects may accordingly be more inclined to reveal in Unaware.

Hypothesis 4. *Revelation rates in Entitlement will be higher than those in Sequential.*

Hypothesis 5. *Revelation rates in Unaware will be higher than those in Sequential.*

6 Results

Our data were generated in groups of six participants. For each treatment, we have observations from 11 (Baseline and Unaware) or 12 (Sequential and Entitlement) groups. As our experiments are one shot, this yields for each treatment 66 or 72 reveal/conceal decisions. Since observations within groups may not be fully independent, when running non-parametric tests, we reduce the data obtained in a group to one single observation.

6.1 Coordination on equilibria

The left panel of Figure 2 shows whether groups coordinate on an equilibrium. It includes the standard *Nash equilibrium* (dark blue bars, M1), other *monotonic* profiles (medium blue, M2-M6) and *non-monotonic* (light blue) outcomes.

[14] The calibrated FS model in Table 1 in Appendix B indicates that the conditions on M6 are satisfied in about 55.7% of all cases.

[15] A number of economics experiments have shown that winning a real-effort contest can induce entitlement. Examples include Fahr and Irlenbusch (2000), Konow (2000), Oxoby and Spraggon (2008), and Fischer and Normann (2019).

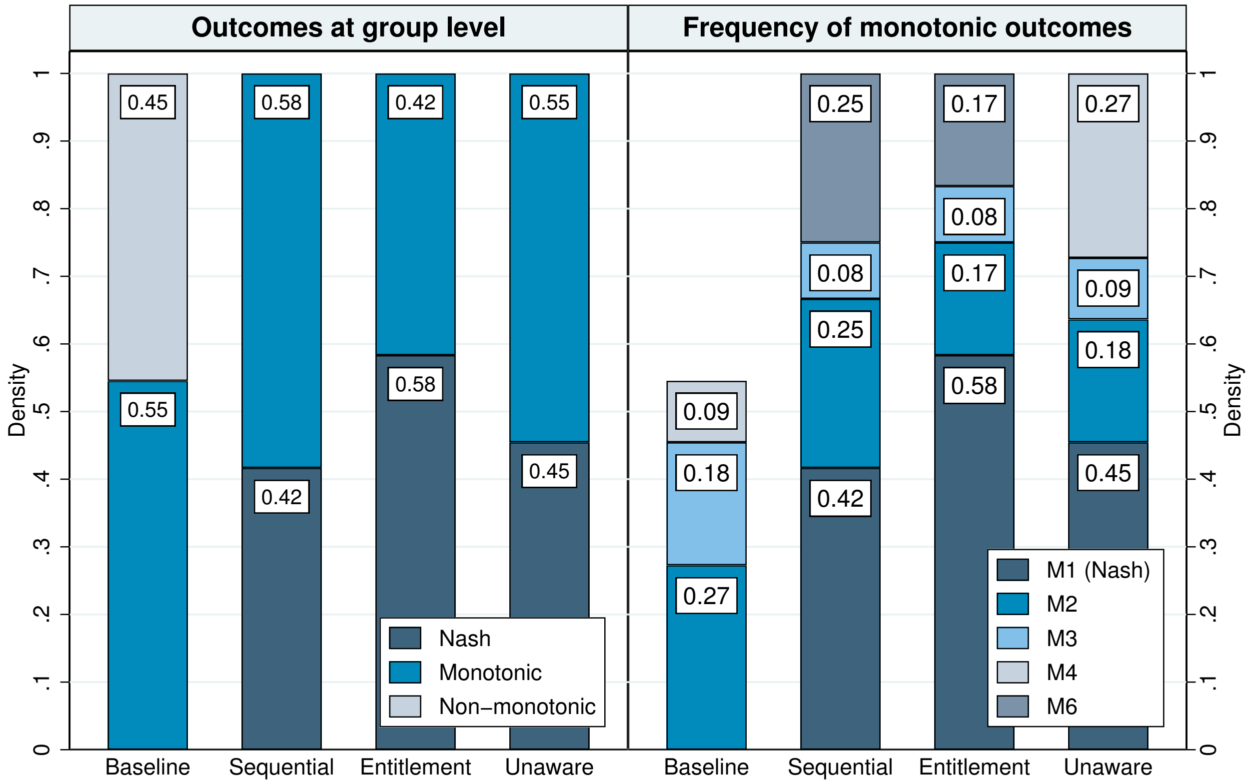

Figure 2: The left panel shows the share of groups who coordinate on monotonic and non-monotonic outcomes. The right panel provides further information on which monotonic outcomes are reached.

A first observation is that coordination failure occurs in Baseline only. Actually, almost half (45%) of the groups do not coordinate on an equilibrium (M1-M6).[16] Our three treatments, by contrast, do not have any coordination failure at all. Statistical support for our Hypothesis 1 can be obtained by counting the groups that coordinate/do not coordinate and comparing these counts with two-sided Fisher-exact tests. A 4×2 test with all treatments is highly significant ($p < 0.001$), as are all pairwise comparisons of Baseline and the treatments with quasi-sequential decision-making are significant at $p = 0.0137$ (Sequential, Entitlement) or $p = 0.0351$ (Unaware), rejecting the null hypothesis.

Result 1. *Groups in Sequential, Entitlement and Unaware coordinate on monotonic outcomes (equilibria) significantly more often than groups in Baseline, supporting Hypothesis 1.*

A second observation immediate in Figure 2 (left panel) is that the standard Nash equilibrium (M1) does not occur in Baseline at all but frequently so in the treatments (42–58%). That is, removing the level-k requirements often leads to full information unraveling. Statistical support comes, firstly, from a two-sided 4×2 Fisher-exact test with all treatments ($p = 0.015$). Secondly, two-sided 2×2 Fisher-exact tests comparing Baseline and our treatments range from $p = 0.0046$ (Entitlement) to $p = 0.0373$ (Sequential). This is consistent with our Hypothesis 2.

Result 2. *Groups in Sequential, Entitlement, and Unaware coordinate on the standard Nash equilibrium with full revelation significantly more often than groups in Baseline, supporting Hypothesis 2.*

[16]This value (45%) is remarkably close to the one our calibration exercise in Section 3.3.3 and Figure 1 suggests, namely 46.7%.

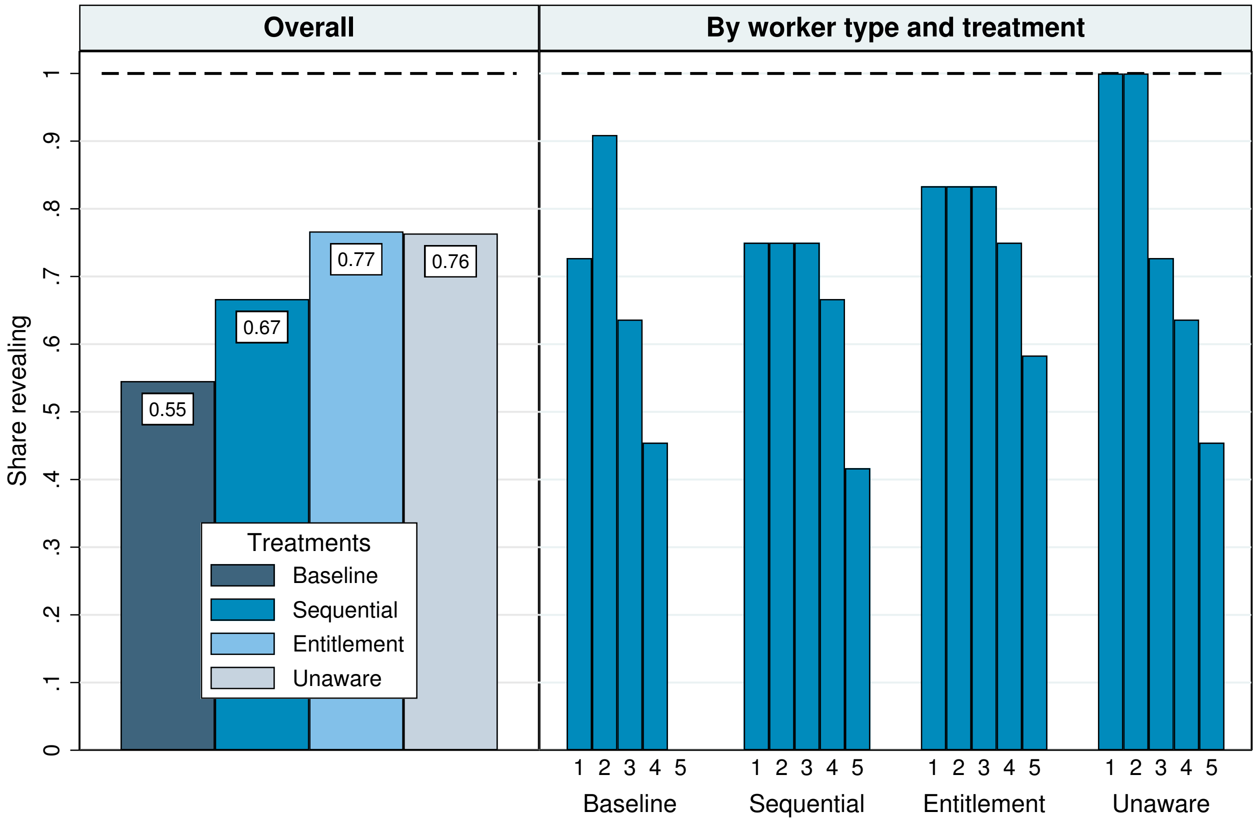

Figure 3: The left panel shows the average revelation rates across our four treatments and the right panel shows the same metric by worker type and treatment. The dashed lines indicate equilibrium predictions. Note that workers 6 (who should – and indeed do – not reveal) are excluded.

The right panel of Figure 2 further disentangles on which monotonic outcomes players coordinate. In Baseline, subjects coordinate on the outcomes M2, M3, and M4 which are also the most likely outcomes according to the level-k analysis (compare Figure 1).[17] In the other treatments, subjects coordinate on M1 (the static Nash equilibrium) most of the time. All players concealing (M6) occurs in Sequential and Entitlement only.

6.2 Revelation rates

In what follows, we exclude the decisions of the type-6 workers. The reason is that concealing is a dominant strategy for worker 6. Since each and every one of our worker-6 subjects did actually conceal, their decision data does not yield any insights regarding our treatments.

The overall revelation rates across treatments are summarized in the left panel of Figure 3. There are substantial differences across treatments. The share of subjects who should reveal in equilibrium and choose to reveal ranges from about 55% in Baseline to about 77% in Entitlement and Unaware, an increase of 42%. Revelation rates in Sequential are intermediate and differ by roughly 10 percentage points to Baseline on the one hand and Entitlement and Unaware on the other.

In Section 5, we argue that removing strategic uncertainty might result in more or less unraveling depending on whether or not eliminating the cognitive requirements for low-productivity workers outweighs the elimination

[17] The ranking is different though. Figure 1 in the Appendix suggests that M3 should be more frequent than M4 which, in turn, should be more frequent than M2. What we observe instead is that M2 is more frequent than M3 which is more frequent than M4.

of coordination issues. The experimental data seems to support the argument for more unraveling. The share of subjects who reveal increases from about 55% in Baseline to about 67% in Sequential. This difference is insignificant (MWU, $p = 0.139$, two sided), but it is evidence showing that there is not less unraveling in Sequential compared to Baseline.

Result 3. *Revelation rates tend to be higher in Sequential than in Baseline. This is support in favor of Hypothesis 3a and against Hypothesis 3b.*

As for the treatments which are designed to mitigate the influence of social preferences, Entitlement and Unaware, we find no significant differences between either of those treatments and Sequential. Aggregate revelation rates differ insignificantly in a non-parametric test (MWU, both $p > 0.517$, two sided). Hence, we formally reject both hypotheses 4 and 5.

Result 4. *There are no significant differences between the revelation rates in Sequential and Entitlement, so we cannot support Hypothesis 4.*

Result 5. *There are no significant differences between the revelation rates in Sequential and Unaware, so we cannot support Hypothesis 5.*

When we take the joint effect of quasi-sequential decision-making and reducing concerns for inequality into account, we do find significant differences. Both Entitlement and Unaware have higher revelation rates than Baseline (MWU, $p = 0.023$ and $p = 0.055$, two sided).

Result 6. *Revelation rates are significantly higher in Entitlement and Unaware compared to Baseline.*

6.3 Revelation rates across types

The right panel of Figure 3 displays the revelation rates of all worker types in all treatments. Worker types and revelation rates are monotonic in all treatments except Baseline. The distribution of k levels (Arad and Rubinstein, 2012; Benndorf, Kübler, and Normann, 2017) suggest limited revelation should not be expected to be uniform across types. In Baseline, there should be comparatively high revelation rates for workers 1, 2, and 3 (as most subjects meet the level-k requirements for these decisions) whereas workers 4 and 5 (who require rarely-met higher levels of reasoning) should more often fail to reveal only. This is what we observe. The right panel of Figure 3 shows that low-productivity workers such as types 4 or 5 reveal less often in Baseline; in fact, participants in the role of worker 5 literally never reveal.

If limited depth of reasoning is the only behavioral force impeding unraveling, this pattern should disappear for the treatments with quasi-sequential moves — but this is not the case. Calculating the correlation of reveal decisions and worker type, we find negative and significant coefficients: $\rho = -0.709$ for Baseline, $\rho = -0.519$ for Sequential, $\rho = -0.553$ for Entitlement and $\rho = -0.648$ for Unaware. So, the correlations differ and the one for Baseline is larger (in absolute terms), but not significantly so,[18] suggesting that inequality aversion has quite some force.

6.4 Timing of decisions

In a last step, we take a closer look at the dynamics of the decision-making and analyze whether the process is consistent with the hypothetical dynamics behind unraveling. Suppose all workers conceal. In that case, only

[18]The statements on statistical significance in this paragraph are based on linear regressions, clustered at the group level, with *reveal* as the dependent variable and worker *type* as the explanatory variable. For all treatments we find $p < 0.001$ concerning the correlations, and we find $p > 0.1$ when we compare the coefficients across treatments with post-hoc tests.

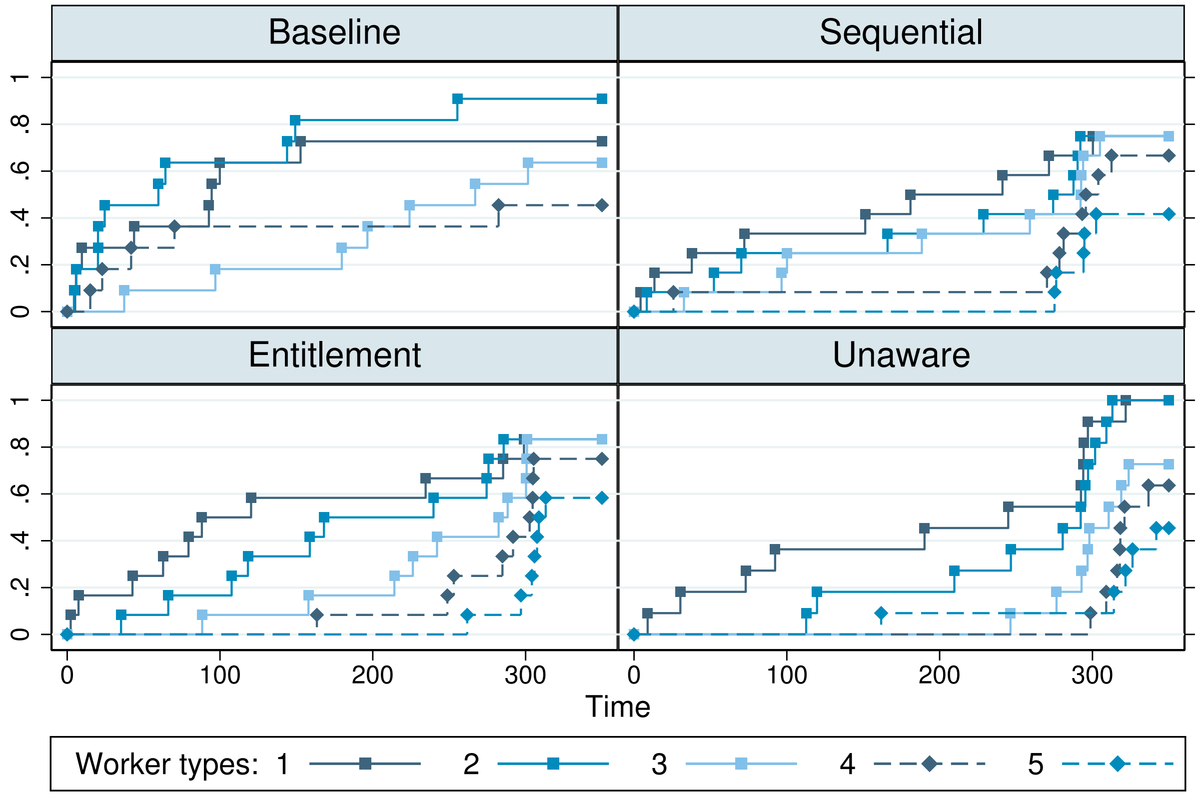

Figure 4: Time of last strategy change. The plots show the share of workers who have ultimately decided to reveal at a given point in time.

worker 1 has an incentive to reveal. Doing so will change the incentives for worker 2 who should then reveal, and so on. Since this logic only applies to workers who reveal, the analyses in this subsection focuses mostly on those workers who reveal at the end of the experiment.

A visual representation of the timing is provided in Figure 4. The horizontal axes show time and the vertical axes document the share of subjects who have already decided that they will reveal. Each line represents one type of the worker, and the graphs are plotted separately for the four treatments. The figure provides some interesting insights. In Sequential, Entitlement, and Unaware, the hypothetical dynamics seem to correctly predict the sequence of decisions. For those workers who reveal at the end of the experiment, type-1 workers reveal first, followed by type-2 workers, followed by type-3 workers, and so on. This is not the case in the Baseline treatment. Here, for example, type-4 workers start revealing before type-3 workers.

Table 3 reports two metrics. The left-hand columns report the average point in time of the last strategy change. Workers who reveal tend to find their decision later, the lower their productivity. Put differently, the relation between the average decision time and the worker type is almost monotonic. The right-hand columns report the share of workers who find their final decision later than the worker with the next higher productivity in their group. It can be seen that this is the case for the majority of the workers in all treatments except Baseline.

Both findings are consistent with the hypothetical thought process mentioned earlier. To formally test whether the patterns are significant we apply non-parametric tests. More precisely, we calculate the differences between the points in time when the workers t and $t-1$ decide to reveal and find that these differences are non-negative in all treatments (two-sided sign tests, all $p < 0.011$) except Baseline ($p = 0.629$). Also note that

	Average time of the last strategy change				Share who decide after worker $t-1$			
Worker	B	S	E	U	B	S	E	U
1	63.1	141.4	122.2	194.5	–	–	–	–
2	74.8	185.5	173.0	252.6	0.43	0.56	0.70	0.64
3	186.2	206.7	240.1	295.5	1.00	0.89	1.00	0.75
4	86.5	257.6	273.3	316.8	0.25	0.75	0.89	0.86
5	–	288.6	299.7	293.1	–	1.00	1.00	0.80

Table 3: Dynamics of the decision-making process for the workers who reveal.

analogous tests for workers who conceal do not reveal such a pattern (all $p > 0.453$).

Result 7. *The timing of subjects' decision-making is consistent with the hypothetical dynamics of unraveling in the treatments with quasi-sequential decision-making.*

As for differences across treatments, we find that the average time of the final choice to reveal or to conceal are 84.9, 195.5, 203.1, and 245.4 seconds into the experiment in Baseline, Sequential, Entitlement, and Unaware, respectively. Two-sided Wilcoxon-Mann-Whitney ranksum tests suggest that subjects find their decision earlier in Baseline compared to all other treatments (MWU, all $p < 0.001$), but there are only few differences between the treatments with quasi-sequential decision-making. Subjects decide later in Unaware compared to Sequential (MWU, $p = 0.037$). The remaining pairwise comparisons (i.e., Entitlement and either Sequential or Unaware) are not statistically significant (both $p > 0.212$).

7 Conclusion

Our paper identifies and disentangles forces that may impede information unraveling. A well-established and influential theoretical literature (Viscusi, 1978; Grossman and Hart, 1980; Grossman, 1981; Milgrom, 1981; Milgrom and Roberts, 1986) shows that information revelation should be complete and immediate. Since "no news is bad news," senders are forced to reveal information. In experiments (Benndorf, Kübler, and Normann, 2015; Jin, Luca, and Martin, 2021; Penczynski and Zhang, 2018; Hagenbach and Perez-Richet, 2018; Benndorf, 2018), information unraveling is often incomplete. Senders do not fully disclose information, and it is not only the players with the least favorable types that decide to conceal. Incomplete unraveling may be due to a limited depth of reasoning[19] (Nagel, 1995; Bosch-Domenech et al., 2002) or other-regarding preferences such as inequality aversion (Fehr and Schmidt, 1999).

Our novel design removes the requirements on subjects' depth of reasoning. As we provide ample decision time and since subjects see the current decisions profile and its payoff implications on the screen, decision-making is quasi-sequential and revealing information does not require great sophistication. At the same time, the new design may facilitate coordination on equilibria with full or partial unraveling which exist with other-regarding preferences. Two of our treatments should reduce other-regarding concerns because they trigger entitlement or make the externality less transparent. Compared to the design of previous experiments, we can thus disentangle the effects of limited depth of reasoning and other-regarding motives: If incomplete revelation persists, it must be due to other-regarding preferences.

Our results show that the new design with quasi-sequential decision-making was successful in purging the requirements on limited depth of reasoning. There are no coordination failures at all in our treatments compared

[19] Already the seminal analysis of Milgrom and Roberts (1986) distinguishes between sophisticated and unsophisticated players. Another model of bounded rationality suggesting limited information unraveling is cursedness (Eyster and Rabin, 2005). See Frondel, Gerster, and Vance (2020) for an application of cursed equilibria on information revelation in the housing market.

to the baseline variant with simultaneous decision-making where nearly half of the groups do not play an equilibrium. Compared to previous experiments, this level of zero coordination failure allows us to more clearly interpret the treatment data with respect to the benchmark of full unraveling.

We find in our new treatments more information unraveling and substantially more groups with complete information revelation. While complete unraveling does not occur at all in the baseline variant, it does occur in 42-58% of the treatment groups – an economically and statistically significant increase. To the best of our knowledge, this is the first time full unraveling occurs in the lab. Average revelation increases by 12 to 22 percentage points in our treatments. This increase of 22 to 40% is statistically significant.[20]

How about monotonic outcomes that are Fehr and Schmidt (1999) equilibria? In the baseline variant, slightly more than half of the groups ended up in such an equilibrium. This share stays roughly constant in our treatments where, again, roughly half of the groups play a monotonic outcome. Since depth of reasoning should no longer play a role, we interpret these monotonic outcomes with incomplete unraveling as FS equilibria. They clearly indicate that other-regarding motives are a significant behavioral explanation for incomplete revelation.

[20]Borderline significant only for treatment Sequential.

References

Ackfeld, Viola and Werner Güth (2019). “Personal information disclosure under competition for benefits: Is sharing caring?” In: *MPI Collective Goods Discussion Paper* 2019/4.

Andersson, Ola and Erik Wengström (2007). “Do antitrust laws facilitate collusion? Experimental evidence on costly communication in duopolies”. In: *Scandinavian Journal of Economics* 109.2, pp. 321–339.

Arad, Ayala and Ariel Rubinstein (2012). “The 11-20 money request game: A level-k reasoning study”. In: *American Economic Review* 102.7, pp. 3561–73.

Balliet, Daniel (2010). “Communication and cooperation in social dilemmas: A meta-analytic review”. In: *Journal of Conflict Resolution* 54.1, pp. 39–57.

Bederson, Benjamin B et al. (2018). “Incomplete disclosure: Evidence of signaling and countersignaling”. In: *American Economic Journal: Microeconomics* 10.1, pp. 41–66.

Benndorf, Volker (2018). “Voluntary disclosure of private information and unraveling in the market for lemons: An experiment”. In: *Games* 9.2, p. 23.

Benndorf, Volker, Dorothea Kübler, and Hans-Theo Normann (2015). “Privacy concerns, voluntary disclosure of information, and unraveling: An experiment”. In: *European Economic Review* 75, pp. 43–59.

— (2017). “Depth of reasoning and information revelation: An experiment on the distribution of k-levels”. In: *International Game Theory Review* 19.04, p. 1750021.

Benndorf, Volker, Holger A Rau, and Christian Sölch (2019a). “Gender differences in motivational crowding out of work performance”. In: *Economic Inquiry* 57.1, pp. 206–226.

— (2019b). “Minimizing learning in repeated real-effort tasks”. In: *Journal of Behavioral and Experimental Finance* 22, pp. 239–248.

Blanco, Mariana, Dirk Engelmann, and Hans Theo Normann (2011). “A within-subject analysis of other-regarding preferences”. In: *Games and Economic Behavior* 72.2, pp. 321–338.

Blume, Andreas and Andreas Ortmann (2007). “The effects of costless pre-play communication: Experimental evidence from games with Pareto-ranked equilibria”. In: *Journal of Economic theory* 132.1, pp. 274–290.

Bochet, Olivier, Talbot Page, and Louis Putterman (2006). “Communication and punishment in voluntary contribution experiments”. In: *Journal of Economic Behavior & Organization* 60.1, pp. 11–26.

Bochet, Olivier and Louis Putterman (2009). “Not just babble: Opening the black box of communication in a voluntary contribution experiment”. In: *European Economic Review* 53.3, pp. 309–326.

Bosch-Domenech, Antoni et al. (2002). “One, two,(three), infinity,...: Newspaper and lab beauty-contest experiments”. In: *American Economic Review* 92.5, pp. 1687–1701.

Brandts, Jordi and David J Cooper (2007). “It’s what you say, not what you pay: An experimental study of manager-employee relationships in overcoming coordination failure”. In: *Journal of the European Economic Association* 5.6, pp. 1223–1268.

Brandts, Jordi and Charles A Holt (1992). “An experimental test of equilibrium dominance in signaling games”. In: *The American Economic Review* 82.5, pp. 1350–1365.

— (1993). “Adjustment patterns and equilibrium selection in experimental signaling games”. In: *International Journal of Game Theory* 22.3, pp. 279–302.

Cadsby, Charles Bram, Murray Frank, and Vojislav Maksimovic (1990). “Pooling, separating, and semiseparating equilibria in financial markets: Some experimental evidence”. In: *The Review of Financial Studies* 3.3, pp. 315–342.

— (1998). “Equilibrium dominance in experimental financial markets”. In: *The Review of Financial Studies* 11.1, pp. 189–232.

Cooper, David J, Susan Garvin, and John H Kagel (1997a). “Adaptive learning vs. equilibrium refinements in an entry limit pricing game”. In: *The Economic Journal* 107.442, pp. 553–575.

— (1997b). “Signalling and adaptive learning in an entry limit pricing game”. In: *The RAND Journal of Economics* 28.4, p. 662.

Cooper, David J and John H Kagel (2003). “Lessons learned: generalizing learning across games”. In: *American Economic Review* 93.2, pp. 202–207.

Crawford, Vincent (1998). “A survey of experiments on communication via cheap talk”. In: *Journal of Economic theory* 78.2, pp. 286–298.

Crawford, Vincent P and Joel Sobel (1982). “Strategic information transmission”. In: *Econometrica: Journal of the Econometric Society*, pp. 1431–1451.

Dickhaut, John W, Kevin A McCabe, and Arijit Mukherji (1995). “An experimental study of strategic information transmission”. In: *Economic Theory* 6.3, pp. 389–403.

Eyster, Erik and Matthew Rabin (2005). “Cursed equilibrium”. In: *Econometrica* 73.5, pp. 1623–1672.

Fahr, René and Bernd Irlenbusch (2000). “Fairness as a constraint on trust in reciprocity: earned property rights in a reciprocal exchange experiment”. In: *Economics Letters* 66.3, pp. 275–282.

Fehr, Ernst and Klaus M Schmidt (1999). “A theory of fairness, competition, and cooperation”. In: *The quarterly journal of economics* 114.3, pp. 817–868.

Fischbacher, Urs (2007). “z-Tree: Zurich toolbox for ready-made economic experiments”. In: *Experimental economics* 10.2, pp. 171–178.

Fischer, Christian and Hans-Theo Normann (2019). “Collusion and bargaining in asymmetric Cournot duopoly—An experiment”. In: *European Economic Review* 111, pp. 360–379.

Fonseca, Miguel A and Hans-Theo Normann (2012). “Explicit vs. tacit collusion—The impact of communication in oligopoly experiments”. In: *European Economic Review* 56.8, pp. 1759–1772.

Forsythe, R., M. R. Isaac, and T. R. Palfrey (1989). “Theories and tests of “blind bidding” in sealed-bid auctions”. In: *RAND Journal of Economics* 20.2, pp. 214–238.

Frondel, Manuel, Andreas Gerster, and Colin Vance (2020). “The power of mandatory quality disclosure: Evidence from the German housing market”. In: *Journal of the Association of Environmental and Resource Economists* 7.1, pp. 181–208.

Gächter, Simon et al. (2009). “Are experimental economists prone to framing effects? A natural field experiment”. In: *Journal of Economic Behavior & Organization* 70.3, pp. 443–446.

Greiner, Ben (2015). “Subject pool recruitment procedures: organizing experiments with ORSEE”. In: *Journal of the Economic Science Association* 1.1, pp. 114–125.

Grossman, Sanford J (1981). “The informational role of warranties and private disclosure about product quality”. In: *Journal of Law and Economics* 24.3, pp. 461–483.

Grossman, Sanford J and Oliver D Hart (1980). “Disclosure laws and takeover bids”. In: *Journal of Finance* 35.2, pp. 323–334.

Güth, Werner et al. (2019). “Compulsory Disclosure of Private Information: Theoretical and Experimental Results for the Acquiring-a-Company Game”. In: *Journal of Institutional and Theoretical Economics* 175.3, pp. 502–523.

Hagenbach, Jeanne and Eduardo Perez-Richet (2018). “Communication with Evidence in the Lab”. In: *Games and Economic Behavior* 112, pp. 139–165.

Heiny, Friederike, Tianchi Li, and Michel Tolksdorf (2020). “We Value Your Privacy: Behavior-Based Pricing Under Endogenous Privacy”. In: *Available at SSRN 3508762*.

Isaac, R Mark and James M Walker (1988). “Communication and free-riding behavior: The voluntary contribution mechanism”. In: *Economic inquiry* 26.4, pp. 585–608.

Jin, Ginger Zhe, Michael Luca, and Daniel Martin (2021). "Is no news (perceived as) bad news? An experimental investigation of information disclosure". In: *American Economic Journal: Microeconomics* 13.2, pp. 141–73.

Konow, James (2000). "Fair shares: Accountability and cognitive dissonance in allocation decisions". In: *American economic review* 90.4, pp. 1072–1091.

Kübler, Dorothea, Wieland Müller, and Hans-Theo Normann (2008). "Job-market signaling and screening: An experimental comparison". In: *Games and Economic Behavior* 64.1, pp. 219–236.

Luca, Michael and Jonathan Smith (2013). "Salience in quality disclosure: evidence from the US News college rankings". In: *Journal of Economics & Management Strategy* 22.1, pp. 58–77.

Mathios, Alan D (2000). "The impact of mandatory disclosure laws on product choices: An analysis of the salad dressing market". In: *The Journal of Law and Economics* 43.2, pp. 651–678.

Milgrom, Paul and John Roberts (1986). "Relying on the information of interested parties". In: *RAND Journal of Economics* 17.1, pp. 18–32.

Milgrom, P.R. (1981). "Good news and bad news: Representation theorems and applications". In: *Bell Journal of Economics* 12.2, pp. 380–391.

Miller, Ross M and Charles R Plott (1985). "Product quality signaling in experimental markets". In: *Econometrica: Journal of the Econometric Society*, pp. 837–872.

Moellers, Claudia, Hans-Theo Normann, and Christopher M Snyder (2017). "Communication in vertical markets: Experimental evidence". In: *International Journal of Industrial Organization* 50, pp. 214–258.

Nagel, Rosemarie (1995). "Unraveling in guessing games: An experimental study". In: *The American Economic Review* 85.5, pp. 1313–1326.

Oprea, Ryan, Gary Charness, and Daniel Friedman (2014). "Continuous time and communication in a public-goods experiment". In: *Journal of Economic Behavior & Organization* 108, pp. 212–223.

Oxoby, Robert J and John Spraggon (2008). "Mine and yours: Property rights in dictator games". In: *Journal of Economic Behavior & Organization* 65.3-4, pp. 703–713.

Penczynski, Stefan P and Sihong Zhang (2018). *Disclosure of Verifiable Information under Competition: An Experimental Study.* Tech. rep. Mimeo.

Potters, Jan and Frans Van Winden (1996). "Comparative statics of a signaling game: An experimental study". In: *International Journal of Game Theory* 25.3, pp. 329–353.

Spence, Michael (1973). "Job Market Signaling". In: *The Quarterly Journal of Economics* 87.3, pp. 355–374.

Viscusi, W. K. (1978). "A Note on "Lemons" Markets with Quality Certification". In: *Bell Journal of Economics* 9.1, pp. 277–279.

A Additional Figure

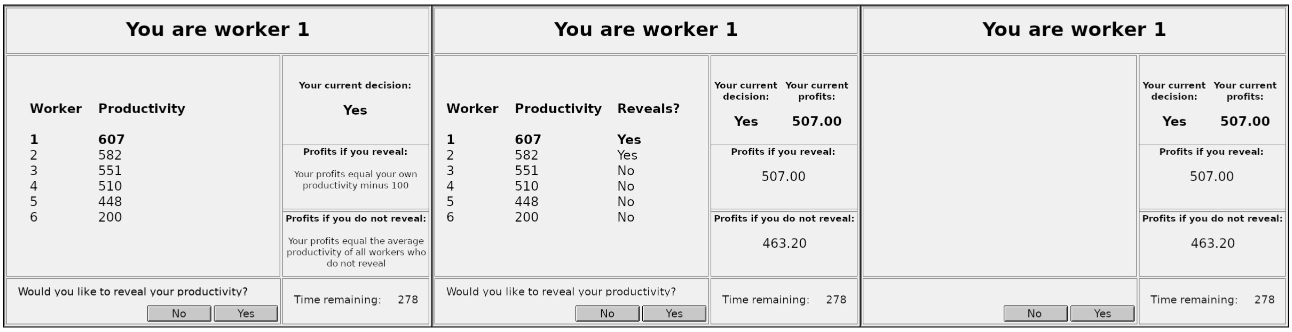

Figure 5: Reproduction of the decision screens in all treatments. The left panel is for Baseline, the middle panel for Sequential and Entitlement, and the right panel is for Unaware.

B Proof of the proposition

We prove the proposition in the main text in three steps. Lemma 1 establishes that all monotonic strategy profiles may be FS equilibria for some FS parameters and we list the conditions on those parameters in Table 1. Lemma 2 shows that, provided concealing is player i's best reply when i maximizes own payoffs, then concealing is also player i's best reply with FS preferences. As Lemma 3 shows, Lemma 2 implies that there are no other FS equilibria.

Lemma 1. *For the parameterization of the experiment, all monotonic strategy profiles may be FS equilibria.*

Proof. From Table 4, there are six monotonic strategy profiles, M1 to M6. All monotonic outcomes have $I_6 = 0$ because it is worker 6's dominant action. Whether these monotonic outcomes are FS equilibria depends on the α and β parameters of the players. Table 4 lists all possible trade-offs for all workers in all equilibria. The requirements for some strategy combination to be an FS equilibrium are referred to as "Condition 1," and "Condition 2" simplifies "Condition 1." For worker 6, concealing is a dominant action also with FS preferences because deviating to revelation decreases their own payoff and increases average disadvantageous inequality. The remaining workers may either reveal or conceal, depending on their FS parameters. $\square$

Lemma 2. *Suppose concealing is player i's best reply with standard preferences. Then concealing is also player i's best reply with FS preferences.*

Proof. Let x_i and x_i' denote player i's monetary payoffs when player i conceals and reveals, respectively, given any action profile for players $j \neq i$. The lemma posits that $x_i \geq x_i'$, so concealing is player i's best reply with standard preferences.

Define the set of all players $j \neq i$ who conceal as $C = \{j \neq i | I_j = 0\}$ and the set of all players $j \neq i$ who reveal as $R = \{j \neq i | I_j = 1\}$. Using C and R, we can define the following shortcut notation. When i conceals, the aggregate *disadvantageous* inequality with respect to the players $\neq i$ who conceal can be defined as $\Delta_C = \sum_{j \in C} \max\{x_j - x_i, 0\}|_{I_i=0}$, and the aggregate *disadvantageous* inequality with respect to the players $\neq i$ who reveal is $\Delta_R = \sum_{j \in R} \max\{x_j - x_i, 0\}|_{I_i=0}$. Still, considering i conceals, the aggregate *advantageous* inequality with respect to the players $\neq i$ who conceal and reveal reads $\Gamma_C = \sum_{j \in C} \max\{x_i - x_j, 0\}|_{I_i=0}$ and $\Gamma_R = \sum_{j \in R} \max\{x_i - x_j, 0\}|_{I_i=0}$, respectively. We use prime ($'$) notation to define analogously the notation for the cases where i reveals ($I_i = 1$), namely Δ_C', Δ_R', Γ_C' and Γ_R'.

We can now compare players i's FS utility as in (2) for i's two actions to determine i's best reply with FS preferences. Concealing yields a (weakly) higher FS utility than revealing if and only if

$$x_i - \frac{\alpha_i}{n-1}(\Delta_C + \Delta_R) - \frac{\beta_i}{n-1}(\Gamma_C + \Gamma_R) \;\geq\; x'_i - \frac{\alpha_i}{n-1}(\Delta'_C + \Delta'_R) - \frac{\beta_i}{n-1}(\Gamma'_C + \Gamma'_R)$$

Note here that $\Delta_C = \Gamma_C = 0$ because all players who conceal get the same monetary payoff. Further, we can set $\Delta'_C = \Gamma'_C = 0$ without loss of generality: If the inequality is met for $\Delta'_C = \Gamma'_C = 0$, then it holds *a fortiori* if $\Delta'_C > 0$ or $\Gamma'_C > 0$. Rearranging, concealing yields a higher FS utility than revealing when

$$(n-1)(x_i - x'_i) \geq \alpha_i(\Delta_R - \Delta'_R) + \beta_i(\Gamma_R - \Gamma'_R)$$

The payoffs of players $j \in R$ are the same, regardless of whether player i reveals or conceals. But since $x_i \geq x'_i$ as assumed in the lemma, revealing (compared to concealing) results in greater or equal disadvantageous inequality, so $\Delta_R - \Delta'_R \leq 0$, and lower or equal advantageous inequality, so $\Gamma_R - \Gamma'_R \geq 0$. (Note that the difference of disadvantageous (advantageous) inequality is zero only if player i is the player with the highest (lowest) payoff for both actions.) From $\Delta_R - \Delta'_R \leq 0$, we obtain $\beta_i(\Gamma_R - \Gamma'_R) \geq \alpha_i(\Delta_R - \Delta'_R) + \beta_i(\Gamma_R - \Gamma'_R)$. The difference of advantageous inequality with respect to a single revealing player is $x_i - x'_i$, so the aggregate difference cannot exceed $(n-1)(x_i - x'_i)$. This implies $(n-1)(x_i - x'_i) \geq \Gamma_R - \Gamma'_R$. To prove the proposition, it is thus sufficient to show

$$(n-1)(x_i - x'_i) \geq \beta_i(n-1)(x_i - x'_i)$$

This holds since $x_i - x'_i \geq 0$ and $\beta_i < 1$ strictly. □

Lemma 3. *There are no FS equilibria in pure strategies other than the monotonic outcomes in Table 1.*

Proof. It is straightforward to see that, for the parameterization used in the experiment ($n = 6$, $c = 100$, and $\Theta = \{607, 582, 551, 510, 448, 200\}$), worker i does not have an incentive (with standard preferences) to reveal unless all other workers with higher productivities $j < i$ reveal. This implies that any strategy profile not listed in Table 4 includes at least one worker who reveals even though concealing would maximize own monetary payoffs. From Proposition 2, such strategy profiles cannot be FS equilibria (in pure strategies). □

Worker	Reveals	Condition 1	Condition 2
1	1	$507 - 128.8 \cdot \beta_1 \geq 403.5 - 26.5 \cdot \alpha_1 - 11.1 \cdot \beta_1$	$26.5 \cdot \alpha_1 \geq -103.5 + 117.7 \cdot \beta_1$
2	1	$482 - 5 \cdot \alpha_2 - 103.8 \cdot \beta_2 \geq 391 - 39 \cdot \alpha_2 - 8.6 \cdot \beta_2$	$34 \cdot \alpha_2 \geq -91 + 95.2 \cdot \beta_2$
3	1	$451 - 17.4 \cdot \alpha_3 - 79 \cdot \beta_3 \geq 375.5 - 54.5 \cdot \alpha_3 - 5.5 \cdot \beta_3$	$37.1 \cdot \alpha_3 \geq -75.5 + 73.5 \cdot \beta_3$
4	1	$410 - 42 \cdot \alpha_4 - 54.4 \cdot \beta_4 \geq 355 - 75 \cdot \alpha_4 - 1.4 \cdot \beta_4$	$33 \cdot \alpha_4 \geq -55 + 53 \cdot \beta_4$
5	1	$348 - 91.6 \cdot \alpha_5 - 29.6 \cdot \beta_5 \geq 324 - 110.8 \cdot \alpha_5$	$19.2 \cdot \alpha_5 \geq -24 + 29.6 \cdot \beta_5$
6	0	$200 - 239.6 \cdot \alpha_6 \geq 100 - 339.6 \cdot \alpha_6$	$100 \cdot \alpha_6 \geq -100$

The standard Nash equilibrium (M1) is a Fehr-Schmidt equilibrium in 90.5% of all cases.

Worker	Reveals	Condition 1	Condition 2
1	1	$507 - 108.80 \cdot \beta_1 \geq 418.33 - 19.27 \cdot \alpha_1 - 1.67 \cdot \beta_1$	$19.27 \cdot \alpha_1 \geq -88.67 + 107.13 \cdot \beta_1$
2	1	$482 - 5 \cdot \alpha_2 - 83.80 \cdot \beta_2 \geq 410 - 27.60 \cdot \alpha_2$	$22.60 \cdot \alpha_2 \geq -72 + 83.80 \cdot \beta_2$
3	1	$451 - 17.40 \cdot \alpha_3 - 59 \cdot \beta_3 \geq 399.67 - 40 \cdot \alpha_3$	$22.60 \cdot \alpha_3 \geq -51.33 + 59 \cdot \beta_3$
4	1	$410 - 42 \cdot \alpha_4 - 34.40 \cdot \beta_4 \geq 386 - 56.40 \cdot \alpha_4$	$14.40 \cdot \alpha_4 \geq -24 + 34.40 \cdot \beta_4$
5	0	$324 - 110.80 \cdot \alpha_5 \geq 348 - 91.60 \cdot \alpha_5 - 29.60 \cdot \beta_5$	$-19.20 \cdot \alpha_5 \geq 24 - 29.60 \cdot \beta_5$
6	0	$324 - 110.80 \cdot \alpha_6 \geq 100 - 359.60 \cdot \alpha_6$	$248.80 \cdot \alpha_6 \geq -224$

M2 is a Fehr-Schmidt equilibrium in 1.2% of all cases.

Worker	Reveals	Condition 1	Condition 2
1	1	$507 - 88.80 \cdot \beta_1 \geq 441.25 - 10.10 \cdot \alpha_1$	$10.10 \cdot \alpha_1 \geq -65.75 + 88.80 \cdot \beta_1$
2	1	$482 - 5 \cdot \alpha_2 - 63.80 \cdot \beta_2 \geq 435 - 17.60 \cdot \alpha_2$	$12.60 \cdot \alpha_2 \geq -47 + 63.80 \cdot \beta_2$
3	1	$451 - 17.40 \cdot \alpha_3 - 39 \cdot \beta_3 \geq 427.25 - 26.90 \cdot \alpha_3$	$9.50 \cdot \alpha_3 \geq -23.75 + 39 \cdot \beta_3$
4	0	$386 - 56.40 \cdot \alpha_4 \geq 410 - 42 \cdot \alpha_4 - 34.40 \cdot \beta_4$	$-14.40 \cdot \alpha_4 \geq 24 - 34.40 \cdot \beta_4$
5	0	$386 - 56.40 \cdot \alpha_5 \geq 348 - 82 \cdot \alpha_5$	$25.60 \cdot \alpha_5 \geq -38$
6	0	$386 - 56.40 \cdot \alpha_6 \geq 100 - 379.60 \cdot \alpha_6$	$323.20 \cdot \alpha_6 \geq -286$

M3 is a Fehr-Schmidt equilibrium in 7.2% of all cases.

Worker	Reveals	Condition 1	Condition 2
1	1	$507 - 68.80 \cdot \beta_1 \geq 463.20 - 3.76 \cdot \alpha_1$	$3.76 \cdot \alpha_1 \geq -43.80 + 68.80 \cdot \beta_1$
2	1	$482 - 5 \cdot \alpha_2 - 43.80 \cdot \beta_2 \geq 458.20 - 9.76 \cdot \alpha_2$	$4.76 \cdot \alpha_2 \geq -23.80 + 43.80 \cdot \beta_2$
3	0	$427.25 - 26.90 \cdot \alpha_3 \geq 451 - 17.40 \cdot \alpha_3 - 39 \cdot \beta_3$	$-9.50 \cdot \alpha_3 \geq 23.75 - 39 \cdot \beta_3$
4	0	$427.25 - 26.90 \cdot \alpha_4 \geq 410 - 33.80 \cdot \alpha_4 - 6.20 \cdot \beta_4$	$6.90 \cdot \alpha_4 \geq -17.25 - 6.20 \cdot \beta_4$
5	0	$427.25 - 26.90 \cdot \alpha_5 \geq 348 - 102 \cdot \alpha_5$	$75.10 \cdot \alpha_5 \geq -79.25$
6	0	$427.25 - 26.90 \cdot \alpha_6 \geq 100 - 399.60 \cdot \alpha_6$	$372.70 \cdot \alpha_6 \geq -327.25$

M4 is a Fehr-Schmidt equilibrium in 9.5% of all cases.

Worker	Reveals	Condition 1	Condition 2
1	1	$507 - 48.80 \cdot \beta_1 \geq 483$	$0 \geq -24 + 48.80 \cdot \beta_1$
2	0	$458.20 - 9.76 \cdot \alpha_2 \geq 482 - 5 \cdot \alpha_2 - 43.80 \cdot \beta_2$	$-4.76 \cdot \alpha_2 \geq 23.80 - 43.80 \cdot \beta_2$
3	0	$458.20 - 9.76 \cdot \alpha_3 \geq 451 - 11.20 \cdot \alpha_3 - 12.80 \cdot \beta_3$	$1.44 \cdot \alpha_3 \geq -7.20 - 12.80 \cdot \beta_3$
4	0	$458.20 - 9.76 \cdot \alpha_4 \geq 410 - 47.60 \cdot \alpha_4$	$37.84 \cdot \alpha_4 \geq -48.20$
5	0	$458.20 - 9.76 \cdot \alpha_5 \geq 348 - 122 \cdot \alpha_5$	$112.24 \cdot \alpha_5 \geq -110.20$
6	0	$458.20 - 9.76 \cdot \alpha_6 \geq 100 - 419.60 \cdot \alpha_6$	$409.84 \cdot \alpha_6 \geq -358.20$

M5 is a Fehr-Schmidt equilibrium in 14.5% of all cases.

Worker	Reveals	Condition 1	Condition 2
1	0	$483 \geq 507 - 48.80 \cdot \beta_1$	$0 \geq 24 - 48.80 \cdot \beta_1$
2	0	$483 \geq 482 - 18.80 \cdot \beta_2$	$0 \geq -1 - 18.80 \cdot \beta_2$
3	0	$483 \geq 451 - 18.40 \cdot \alpha_3$	$18.40 \cdot \alpha_3 \geq -32$
4	0	$483 \geq 410 - 67.60 \cdot \alpha_4$	$67.60 \cdot \alpha_4 \geq -73$
5	0	$483 \geq 348 - 142 \cdot \alpha_5$	$142 \cdot \alpha_5 \geq -135$
6	0	$483 \geq 100 - 439.60 \cdot \alpha_6$	$439.60 \cdot \alpha_6 \geq -383$

M6 is a Fehr-Schmidt equilibrium in 55.7% of all cases.

Table 4: Fehr-Schmidt equilibria in pure strategies. Condition 1 compares the FS utility when sticking to the equilibrium strategy (left-hand side) to the utility from deviation (right-hand side), and Condition 2 is a simplified version of Condition 1. The equilibrium frequencies are based on Blanco, Engelmann, and Normann's (2011) joint distribution of FS parameters.