

Fehrler, Sebastian; Hughes, Niall

Conference Paper

How Transparency Kills Information Aggregation (And Why That May Be A Good Thing)

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik
- Session: Experimental Economics I, No. A09-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Fehrler, Sebastian; Hughes, Niall (2014) : How Transparency Kills Information Aggregation (And Why That May Be A Good Thing), Beiträge zur Jahrestagung des Vereins für Socialpolitik 2014: Evidenzbasierte Wirtschaftspolitik - Session: Experimental Economics I, No. A09-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg

This Version is available at:

<https://hdl.handle.net/10419/100440>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

How Transparency Kills Information Aggregation (And Why That May Be A Good Thing)*

Sebastian Fehrler

Niall Hughes

University of Zurich, and IZA

University of Warwick

sebastian.fehrler@uzh.ch

N.E.Hughes@warwick.ac.uk

November 2013

Abstract

We show theoretical and experimental results that demonstrate the potential of transparency to influence committee decision making and deliberation. We present a model in which committee members have career concerns and unanimity is needed to change the status quo. We study three scenarios - *secrecy*, where votes and communication are secret, *mild transparency*, where individual votes are public, and *full transparency*, where both communication and individual votes are public. The two transparency regimes affect deliberation in different ways but lead both to a breakdown of information aggregation. However, transparency lowers the probability of an error if maintaining the status quo is the correct choice. As a consequence, the principal is better off in expectation under transparency than under secrecy if the cost of wrongly changing the status quo is high enough. We test the model's key predictions in a laboratory experiment varying the level of transparency between treatments. We observe strong effects of transparency on committee error rates that are largely consistent with the model. On the individual level, we observe strong effects on deliberative behavior which go in the predicted direction but are less pronounced than in theory.

Keywords: Committee Decision Making, Deliberation, Transparency, Career Concerns, Information Aggregation, Experiments, Voting, Strategic Communication.

JEL Classification Numbers: C92, D71, D72.

*We thank Alessandra Casella, Aniol Llorente-Saguer, Alessandro Lizzeri, Massimo Morelli, Becky Morton, Maik Schneider, Jean-Robert Tyran and participants at several conferences and seminars for their valuable comments.

1. Introduction

“I’ve searched all the parks in all the cities and found no statues of committees.” (G.K. Chesterton)

The quote above illustrates nicely how we are often willing to delegate decision making to committees, yet praise and recognition remain firmly attached to the individual. Moreover, promotions in organizations are usually not granted to groups but to individuals. Many committees are therefore composed of members who seek recognition from the principal in order to advance in their careers. Obvious examples are committees in corporations deciding about business strategies, campaign committees of political parties and even search committees in academia.

Recent research has shown that career concerns of a single agent can have negative consequences for a principal (Holmström (1999); Prat (2005); Fox and Van Weelden (2012)). While the principal would like the agent’s action to maximize her expected utility, the agent behaves so as to make himself appear smart to the principal. These two objectives don’t necessarily align. Further research shows that career concerns might also be of great importance for committee decision making (Levy (2007); Visser and Swank (2007, 2012); Meade and Stasavage (2008); Gersbach and Hahn (2008)). All of these papers study career concerns under different levels of transparency and suggest that the degree of transparency under which a committee deliberates and decides plays an important role.

In practice, different transparency rules have been adopted by such important committees as the Federal Reserve Open Market Committee (FOMC) and the Governing Council of the European Central Bank. While the latter only announces the Council’s decisions the FOMC publishes individual voting records and since 1993 also minutes of its meetings (Meade and Stasavage (2008)). As it is certainly also true for many other committees that their decision rules vary with respect to the level of transparency, it is important to better understand potential effects of such institutional differences.

In this paper we show theoretically and empirically that transparency can have strong effects on deliberation and decision making of committees. We construct a new model

that, in a departure from previous models, allows for largely unstructured communication. This is a key element in decision making of real committees and has been shown to matter a lot when career concerns are not present in a recent experimental study (Goeree and Yariv (2011)). As in most career concerns models, there are two types of committee members (high and low). They, first, receive more or less informative signals about the true state of the world, then have the opportunity to deliberate, and finally vote for or against changing the status quo in favor of an alternative option. Unanimity is required for a change of the status quo. A committee member's utility depends on the principal's belief about the probability that he is a high type.

We study three different transparency regimes - *secrecy*, where votes and communication are secret, *mild transparency*, where individual votes are made public, and *full transparency*, where both communication and individual votes are public. While each regime has more than one equilibrium we argue below that some are more plausible than others. Under secrecy the most plausible equilibrium predicts that committee members pool information and vote unanimously for the state with the higher posterior probability. Under mild transparency better informed committee members do not share their information, in order to distinguish themselves from less well-informed members, and vote according to their signal. Finally, under full transparency all members are predicted to truthfully share their information but still vote according to their own signal in the final vote. This happens because switching would reveal a low quality signal and thereby indicate a low type.

The first key result is that committee members succeed in aggregating information under a secrecy regime but fail to do so under both forms of transparency. The second key result is that the breakdown of information aggregation occurs through different channels. Depending on the level of transparency, committee members either don't pool their information or pool it but disregard it in the final vote. Successful information aggregation results in fewer wrong decisions overall under secrecy. However, we show that when the cost of a mistake in one state of the world far outweighs that in the other state, then transparency serves to improve the principal's utility *vis-a-vis* secrecy - our third key result. Here transparency aids the principal *because* it causes information aggregation to

break down. This makes it very difficult for committee members to come to a unanimous decision. This bias towards the status quo translates into many mistakes in one state of the world but also hardly any in the other. We further show that in this case the principal also prefers unanimity rule to simple majority rule.

To test these predictions, we run a laboratory experiment with groups consisting of two committee members and one principal. To the best of our knowledge this is the first experiment on committee voting with career concerns. As predicted, we find that players largely tell the truth regarding their signal under both secrecy and full transparency. However, subjects lie to a large extent about their type under full transparency and information aggregation is much worse in the final vote. As a result, aggregate error rates are very different under secrecy and full transparency. Secrecy gives rise to far lower error rates when changing the status quo is the right thing to do, but higher error rates if sticking to the status quo is the right choice. This leads to higher payoffs for the principals under transparency than under secrecy, just as predicted by the model. Aggregate level results regarding error rates under mild transparency are similar to the results under full transparency. However, they result from quite different individual level behavior. Mild transparency leads to a significant level of deception regarding the signal, which harms information aggregation, whereas full transparency causes low quality members to stick with their signal even when they think they are wrong. These results confirm our key theoretical predictions and thus demonstrate that transparency can indeed have very strong effects.

We also observe a number of deviations from our theoretical predictions regarding individual level behavior. Interestingly, deliberation regarding the signal under mild transparency and regarding the type under full transparency is more truthful and informative than predicted, even though the truthful subjects pay a price for it in terms of a lower reputation. This suggests that some subjects face psychological costs of lying, and confirms results from previous experimental studies that have also found “too much” truth-telling in deliberation (Cai and Wang (2006); Goeree and Yariv (2011)). However, while overly truthful communication survives under different voting rules in Goeree and Yariv’s (2011) study, we find that the level of truthfulness varies greatly between our treatments, making

transparency appear the more important element of institutional design.

In the next section we present a review of related literature. We then set up and solve the model in full generality. We proceed with the experimental design and theoretical predictions for the chosen parameter values before discussing the aggregate and individual level results. We conclude with a discussion of the main results and future research questions.

2. Related Literature

Two recent papers deal with the issue of transparency in single principal - single agent settings. Prat (2005) shows in a simple setup that it may be better for the principal to observe only the consequences of the agent's choice rather than observing the choice directly. If actions are observed, the agent will have an incentive to disregard his private signal and simply act how a high type is expected to. This leads to too many wrong decisions and makes it more difficult to separate high and low types. Fox and Van Weelden (2012) show that these results are overturned when the prior on one state of the world is sufficiently strong, and the cost of mistakes are asymmetric across states.

Another string of papers have addressed the issue of transparency in the context of committee voting. Levy (2007) examines how the decisions of careerist experts in a committee of three are affected by the degree of voting information that is revealed. In her model, secrecy (where only the group decision is seen) leads members to pander to the *ex ante* most likely state, while transparency (publishing individual voting records) will make reforms more likely to be accepted. Visser and Swank (2007, 2012) present a model in which committee members will vote in the best interest of the principal when communication is public and individual votes are observed. The fact that committee members are made worse off in terms of reputation spurs them into arranging non-transparent pre-meetings. These results rest on the assumption that the state is not revealed to the principal and that players do not know their own types - this drives the committee to present a "united front" regardless of whether it is the right choice or not. Meade and Stasavage (2008) study committee deliberations in a monetary policy setting, in which a known

expert speaks first, followed by two experts of unknown quality. They show that if deliberations are kept private there is a greater likelihood that committee members will dissent (i.e. reveal their signals truthfully) than would be the case if deliberations were public. Gersbach and Hahn (2008) study a two period game in which members of a monetary policy committee must choose an interest rate in each period but each individual may be dismissed after the first period. They find that when communication and voting are public, low types try to masquerade as high types, resulting in many wrong decisions. This behavior allows low types to be weeded out by the second period, but the overall effect of transparency on welfare remains negative.

At this point we would like to highlight the differences between these models and ours, to show why we constructed it, instead of taking one of the existing models to the laboratory. There are a number of key differences: (i) in contrast to Levy (2007) we allow for communication because there are few real world committees without deliberation and we expect this feature to matter; furthermore, the other studies that do allow for deliberation put a rigid structure on it, while we allow for free-form communication after an initial straw poll; (ii) in our model committee members know their own type, which we believe to be a reasonable assumption for many committees, while this is not the case in Visser and Swank (2007, 2012); (iii) we allow agents to change their mind after deliberation and vote for a different alternative than the one they had favoured initially; and (iv) we allow for asymmetric costs of mistakes in the different states of the world, neither of these were allowed for in Gersbach and Hahn (2008); finally, (v) Gersbach and Hahn (2008) compare secrecy with full transparency - where both communication and individual votes are public - while we allow for the intermediate case of mild transparency, where communication is private but votes are not. This form of mild transparency would be consistent with situations where committee members hold secret pre-meetings.

Empirical literature on the role of transparency is quite scarce, as observational data is seldom available and, to the best of our knowledge, no experimental study has been conducted prior to this one.¹ A rare empirical study of career concerned committees is

¹There have been some experimental papers on the standard Holmström (1999) career concerns model; notably Irlenbusch and Sliwka (2006) and Koch et al. (2009).

offered by Meade and Stasavage (2008) who examine the Federal Open Market Committee meeting transcripts before and after a change in transparency. They find that the style of deliberations changed after 1993, when transcripts became public. Consistent with their theoretical predictions, they find that openly expressed dissent is lower under transparency.

A number of important experimental papers have studied other elements of committee decision making, such as the role of the voting rule or the role of deliberation (most of them in a common value set-up and none of them with career concerned committee members). Guarnaschelli et al. (2000) show that while committees come to very different outcomes under different voting rules without deliberation, if a straw poll is allowed players largely aggregate information successfully - over 90% of reports in the straw poll are truthful. Goeree and Yariv (2011) extend this analysis by allowing free-form communication and some heterogeneity of preferences. They show that communication is overwhelmingly truthful - even when players would benefit from lying. Furthermore, the ability to communicate leads committees to the same outcomes regardless of the voting rule used. They conclude that the voting rule might be a less important element of institutional design than deliberation protocols. Our experiment demonstrates that the theoretically predicted outcome differences under different levels of transparency appear very robust to deviations from equilibrium deliberation.

3. The Model

A committee of n members must make decision $D \in \{B(lue), R(ed)\}$ on behalf of a principal. There are two states of the world $S \in \{B, R\}$ with $Pr(S = B) = Pr(S = R) = 0.5$. The utility of the principal in each state is

$$\begin{aligned} U_{principal}(D = B|S = B) &= x \\ U_{principal}(D = R|S = B) &= U_{principal}(D = B|S = R) = 0 \\ U_{principal}(D = R|S = R) &= 1 - x \end{aligned}$$

with $x \in [0.5, 1)$. That is, the principal gets higher utility from a correct group decision when the state is B . The group decision is made by unanimity rule, whereby option R is implemented only if all n members vote for it, otherwise the status quo B is upheld.²

Committee members are not perfectly informed of the state of the world. Instead, each member gets an informative signal about the true state of the world, where the level of informativeness depends on his type. Each member $i \in n$ is either a high or low type, $t_i \in \{H, L\}$, where $Pr(t_i = H) = q$ and $Pr(t_i = L) = 1 - q$. They each receive an informative signal $s_{t_i} \in \{B, R\}$, where the accuracy of the signals are given by $Pr(s_H = S) = \sigma_H$ and $Pr(s_L = S) = \sigma_L$ respectively. High types receive more informative signals than low types, $0.5 < \sigma_L < \sigma_H \leq 1$.

Types and signals are private information. However, once nature has chosen the number of H and L types and these players have received their signals they can communicate with each other. First, in a simple straw poll, each member simultaneously announces a message $m_i \in \{B, R, \emptyset\}$, i.e. he can announce that his signal was in favor of B or R or can remain silent. Next, the committee enters a round of free-form communication where voters can elaborate on their straw poll announcement, reveal their type, discuss which voting strategy is the best for the group etc. Finally, after both rounds of communication, each voter casts a secret ballot $v_i \in \{B, R\}$.³ Once the committee has made its decision, the true state of the world is revealed and the utility of the principal is realized.

Committee members' primary concern is not making the correct decision for the principal but rather to advance their own individual careers or reputations. The payoff of a committee member is simply the principal's posterior belief that he is of type H , given by $\hat{q} \in [0, 1]$. Before the game starts, the principal's prior for each player is $Pr(t_i = H) = q$. She updates this belief as good as she can, where this ability depends on how much of the decision making process she observes, i.e. on the level of transparency. We compare three different transparency regimes: (1) *secrecy*, where the principal only observes the group decision D , (2) *mild transparency*, where the principal also sees how each individual votes, and (3) *full transparency*, where the principal witnesses each player's message

²In section 3.2. we also consider simple majority rule.

³We could also allow for abstention in the final voting stage but it would not change any equilibrium prediction.

m_i , free-form communication, and final vote v_i .

In what follows, we characterize the set of symmetric perfect Bayesian equilibria under the three transparency regimes. As committee members' payoffs depend on the principal's beliefs, there will be many equilibria in each of these settings. We restrict attention to *responsive* equilibria, where each committee member's strategy depends at least on his own signal, and potentially on other player's messages.⁴ We ignore "mirror" equilibria, in which H types vote against what they believe to be the true state. Finally, our equilibria fall into two categories: *Truthful* equilibria, where players believe communication is truthful, and *Non-Truthful* equilibria, where all communication is believed to be babbling.⁵

3.1. Equilibrium

It is useful to first consider how the principal would act if she observed all of the signals directly and made the decision herself. With no career concerns or informational asymmetries to contend with, the principal would simply update her beliefs using Bayes' rule, and maximise her expected utility, choosing option R if the posterior $\hat{P}r(R) > x$, and otherwise B . In a committee of size n let h be the realised number of H types, let j be the number of H types with the correct signal, and let k be the number of L types with the correct signal. Thus, there will be $2j - h$ net high type signals and $2k - n + h$ net low type signals in favour of the true state. Note that each of these numbers may be positive or negative, and while h is observable, j and k are not. Denoting $\#s_H^R \in \{(h - 2j), (2j - h)\}$ as the net number of H type signals observed in favour of R , and $\#s_L^R \in \{(n - h - 2k), (2k - n + h)\}$ as the net number of L type signals for R , the principal will decide in favour of R (B) when

$$\left(\frac{\sigma_H}{1 - \sigma_H}\right)^{\#s_H^R} \left(\frac{\sigma_L}{1 - \sigma_L}\right)^{\#s_L^R} > (<) \frac{x}{(1 - x)} \quad (1)$$

if $\sigma_H < 1$. For $\sigma_H = 1$ the principal will always decide for the state that is indicated

⁴This rules out equilibria in which voters ignore all information, vote randomly, and all gain $\hat{q} = q$.

⁵To illustrate the main results we discuss an example at the end of the next section.

by a high type signal and if there is none for R (B) if

$$\left(\frac{\sigma_L}{1-\sigma_L}\right)^{\#s_L^R} > (<) \frac{x}{(1-x)} \quad (2)$$

However, the principal's decision may be the incorrect one *ex post*. This will happen when the evidence pointing in favor of one option is driven by incorrect rather than correct signals. We can define $k(S)^*$ as the threshold number of correct L type signals below which a wrong decision is taken,

$$k(B)^* = 0.5 \left[n - h + \frac{(h - 2j) \log \left(\frac{\sigma_H}{1-\sigma_H} \right) - \log \left(\frac{x}{1-x} \right)}{\log \left(\frac{\sigma_L}{1-\sigma_L} \right)} \right] \quad (3)$$

$$k(R)^* = 0.5 \left[n - h + \frac{(h - 2j) \log \left(\frac{\sigma_H}{1-\sigma_H} \right) + \log \left(\frac{x}{1-x} \right)}{\log \left(\frac{\sigma_L}{1-\sigma_L} \right)} \right] \quad (4)$$

As these will be real numbers, it is useful to define $\lfloor k(S)^* \rfloor$ as the integer directly before $k(S)^*$, and $I(k(S)^*)$ as an indicator function which equals one if $k(S)^* > 0$ and otherwise equals zero. The probability of the principal making a mistake under the optimal decision rule is then

$$Pr(D \neq S) = \sum_{h=0}^n \binom{n}{h} q^h (1-q)^{n-h} \sum_{j=0}^h \binom{h}{j} \sigma_H^j (1-\sigma_H)^{h-j} \sum_{k=0}^{\lfloor k(S)^* \rfloor} \binom{n-h}{k} I(k(S)^*) \sigma_L^k (1-\sigma_L)^{n-h-k} \quad (5)$$

The decision rule leads to more mistakes in state R than in state B , but because $x > 0.5$, this is optimal.

(1) Secrecy Here the principal can only see the group decision, and so must hold the same posterior \hat{q} for each individual. For this reason committee members have a common interest in making the correct group decision.

Proposition 1. *There always exist truthful equilibria in which each member honestly reveals his signal and type, and the group implements whichever decision has a weighted majority of signals. If the posterior is balanced, the committee randomizes choosing each*

state with equal probability. The probability of a wrong decision in each state is given by (5) with thresholds (3) and (4) where $x = 0.5$ in both cases. The posterior beliefs of the principal after correct and incorrect group decisions are given by:

$$\begin{aligned}\hat{q}_{sec^T}(D \neq S) &= \frac{\sum_{h=0}^n \frac{h}{n} Pr_{sec^T}(D \neq S|h) \binom{n}{h} q^h (1-q)^{n-h}}{Pr_{sec^T}(D \neq S)}, \\ \hat{q}_{sec^T}(D = S) &= \frac{\sum_{h=0}^n \frac{h}{n} (1 - Pr_{sec^T}(D \neq S|h)) \binom{n}{h} q^h (1-q)^{n-h}}{1 - Pr_{sec^T}(D \neq S)}\end{aligned}$$

Proof. See appendix. □

Notice that the probability of a mistake is identical to (5) where $x = 0.5$. This means that, while the principal would like to choose B or R to maximize her expected utility, the committee members instead choose the project which is most likely to align with the state - regardless of how much the principal values it. The formal voting rule plays no role here: all players truthfully reveal their information and then as a group decide to vote unanimously for whichever alternative has a higher posterior probability.

It is intuitive that committee members with a common interest in choosing the correct state will reveal their information and aggregate it optimally. Coughlan (2000) shows that communication between players with a common interest can lead to full aggregation, while Guarnaschelli et al. (2000) and Goeree and Yariv (2011) show that play corresponds overwhelming to truthful equilibria in the laboratory.

(2) Mild Transparency Under this regime, the principal cannot observe any communication, but he does observe the individual votes of committee members as well as the final group decision. Here there can be strong incentives for H types to separate from L types.

Proposition 2. *Both truthful and non-truthful equilibria exist under mild transparency. In any truthful equilibrium the probability of a mistake and the posterior beliefs \hat{q} are the*

same as in truthful equilibria under secrecy. In any non-truthful equilibrium all committee members vote according to their signal and we have

$$\begin{aligned}\hat{q}_{mild^{NT}}(v_i = S) &= \frac{q\sigma_H}{q\sigma_H + (1-q)\sigma_L} \\ \hat{q}_{mild^{NT}}(v_i \neq S) &= \frac{q(1-\sigma_H)}{q(1-\sigma_H) + (1-q)(1-\sigma_L)} \\ Pr_{mild^{NT}}(D = R|S = B) &= (1 - q\sigma_H - (1-q)\sigma_L)^n \\ Pr_{mild^{NT}}(D = B|S = R) &= 1 - (q\sigma_H - (1-q)\sigma_L)^n\end{aligned}$$

Proof. See appendix □

The non-truthful equilibria can come about simply because nobody believes communication to be truthful, or it can be that L types try to communicate their information but H types counteract this by engaging in signal jamming. Note that in the non-truthful equilibria the probability of a mistake in state B is much smaller than the corresponding probability in state R . This occurs because the failure of information aggregation makes it very difficult to get a unanimous decision in favor of R .

Truthful equilibria must be supported by very specific beliefs such that no player has an incentive to deviate. One such set of beliefs is where the principal believes every player to be of low type unless the vote is unanimous.

(3) Full Transparency Under this regime, the principal sees all stages of communication that occur and observes each individual's vote.

Proposition 3. *Both truthful and non-truthful equilibria exist under full transparency. In both cases all committee members vote according to their signal; the probability of a mistake and the posterior beliefs \hat{q} are the same as in non-truthful equilibria under mild transparency. The truthful equilibrium is sustained by out-of-equilibrium beliefs that any members who switches in the final vote from his initial announcement is of type L .⁶*

⁶We discuss different assumptions about out-of-equilibrium beliefs and a slightly different truthful equilibrium in Appendix 1.

Proof. See appendix. □

The non-truthful equilibrium requires that the principal does not take the communication stage into account in his evaluation. However, the equilibrium disappears once committee members believe there is a small probability that messages that correspond to the true state of the world positively influence the principal's evaluation. Therefore, we believe that the truthful equilibrium in which no switching occurs is the most plausible.

While the aggregate outcomes of the truthful and non-truthful equilibria under the transparency regimes coincide, individual behavior in the deliberation stage is very different.

Example Consider the following case: A principal with $x = 0.75$ receives two signals, with potential accuracy $\sigma_H = 1$ and $\sigma_L = 0.55$. With these parameter values the principal will only choose R when there is at least one high type signal for R . When she receives two low type signals for R (or one for R and one for B) the left hand side of equation (2) will be smaller than the right hand side and even though R is the more likely state (equally likely state) the principal will choose B .

Consider now that the signals are not observed by the principal but by members of a two person committee that decides on her behalf and where each member is a high type with probability $q = 0.25$. Assume, first, that the committee operates under secrecy so that the principal can only observe the group decision. Different to what the principal would do if she observed all signals, this committee will choose R in case of two low type signals for R (and mix between voting for R and B in case of conflicting low type signals) which means that it does not behave optimally from the principal's point of view. However, whenever there is a high type signal and a conflicting low type signal the committee optimally aggregates this information and decides for the state that is indicated by the more accurate signal. The committee thereby minimizes the overall probability of a wrong decision. The ex-ante probability of a wrong committee decision is 25.3% in either state of the world.

Assume now that the committee decides under transparency (either mild or full). As both committee members vote according to their signal the ex-ante probability of a wrong

committee decision is much higher if changing the status quo (which requires unanimity) is the right decision (56.1%) than in the other state of the world (11.4%). The reason for the magnitude of this difference is that the probability of conflicting signals in the committee is high.

The overall error rate is 33.8% under transparency and thus higher than under secrecy (25.3%). However, in this example the principal is still better off under transparency because the error rate is lower in the state she cares more about than under secrecy.

3.2. The Optimal Level of Transparency

In this section we assume that the non-truthful equilibrium is played under mild transparency, the truthful under secrecy and either the truthful or the non-truthful under full transparency. The question which regime the principal would favor *ex ante* then reduces to the question whether she prefers secrecy to transparency as she will be indifferent between the two transparency regimes. As transparency and secrecy lead to different error probabilities in the different states of the world, the principal's preference depends on the utilities x and $(1 - x)$ she derives from correct decisions in either state and we can state the following result:

Proposition 4. *There always exists an x^* such that if $x > x^*$ the principal is better off under transparency than under secrecy.*

Proof. See appendix. □

Finally, we can show that the same principals that prefer transparency to secrecy under unanimity rule also prefer unanimity rule with transparency to simple majority rule with either transparency or secrecy - which justifies our focus on the unanimity rule.⁷

⁷For $x = 0.5$ simple majority and for x close to 1 unanimity will be the optimal voting rules, respectively. In general, the optimal voting rule will be some supermajority rule for intermediate levels of x and whether the principal prefers secrecy or transparency under the optimal rule will also depend on x . Further results regarding the optimal decision rule which also depends on the committee size, the signal accuracy and q , can be obtained from the authors upon request.

Proposition 5. (1) Under simple majority rule, the strategies of voters and the beliefs of the principal are the same as under unanimity. Equilibrium error rates will however be different under transparency. (2) For any value of x , the principal prefers secrecy to transparency. (3) A principal with $x > x^*$ prefers unanimity rule with transparency to simple majority rule with either secrecy or transparency.

Proof. See appendix. □

4. Experiment

To test the main predictions of the model, we ran a laboratory experiment with two member committees, where each member was a high type with probability $q = 0.25$, high types received perfectly informative signals ($\sigma_H = 1$) and low types got low accuracy signals ($\sigma_L = 0.55$).⁸ The committee decides under unanimity rule where state B is the status quo. The experiment consists of three treatments - one for each level of transparency.

4.1. Equilibrium Predictions

Ex ante equilibrium error rates are reported in Table 1 for the truthful equilibrium under secrecy, the non-truthful under mild transparency and the truthful or the non-truthful under full transparency. As discussed in section 3, we find these to be the most plausible equilibria. If, instead, subjects under mild transparency played according to the truthful equilibrium, the *ex ante* error rates would be the same as under secrecy.

Table 2 reports the principal's predicted evaluation of player types after observing the true state of the world, the group decision and, in case of transparency, the individual decisions. In the same table we also report the *ex ante* expected evaluations for H and L type voters. Type L voters are predicted to do best under secrecy and H types voters under transparency. The reason is that evaluators are expected to be better able to tell them apart under transparency.

⁸Note that these are exactly the same parameter values as in the example.

Table 1: *Ex ante* Equilibrium Error Rates (in %)

	$S = B$	$S = R$	overall
Secrecy	25.3	25.3	25.3
Mild Transparency	11.4	56.1	33.8
Full Transparency	11.4	56.1	33.8

Note: In the experiment $S = B$ ($S = R$) corresponds to the case of the blue (red) jar (see below).

Table 2: Principal's Equilibrium Beliefs About Members' Types $Pr(t_i = H)$ in %

	Eq. evaluation after correct and wrong decisions		Eq. expected <i>ex ante</i> evaluation for voters	
	dec. corr.	dec. wr.	type H	type L
Secrecy (group decision)	33.5	0	33.5	22.2
Mild Transparency (individual decision)	37.7	0	37.7	20.8
Full Transparency (individual decision)	37.7	0	37.7	20.8

Note: In equilibrium under secrecy the evaluator only takes the group decision into account for her evaluation, under mild transparency he exclusively looks at individual decisions, and under full transparency at messages and individual decisions (which are the same).

Voters should succeed in aggregating information under secrecy but fail completely under the two transparency regimes. Under secrecy all voters are expected to truthfully announce their signal in the straw poll and if needed their type in free-form communication. Under mild transparency we expect communication to be non-truthful, particularly from H types - that is we expect the announcement of signals and types to hold no useful information. Finally, under full information we predict that all player types truthfully announce their signals in the straw poll. However, in free-form communication there will be no new information, as each player has an incentive to announce himself as a H type. So, under mild transparency information aggregation fails because H types withhold information, while mistakes arise under full transparency because L types prefer to stick to their signal announcement, rather than revealing themselves as a L type by switching to

the more likely option.

4.2. Experimental Design

We ran six sessions, two for each transparency regime (see Table 3). Each session consisted of 20 rounds with random matching of subjects into groups. In the first round of the experiment, subjects were randomly assigned to matching groups of nine people.⁹ In every period, new groups of three were randomly formed within the matching groups. This was done to avoid the emergence of reciprocal behavior and at the same time provide independent matching groups. In each group and round, one member was randomly assigned the role of the principal (called the “observer” in the instructions) and the other two were assigned the role of committee members (called “voters” in the instructions). With probability $q = 0.25$ a voter was of type H (“well-informed voter”), with probability $1 - q = 0.75$ of type L (“informed voter”). The task of the voters was to vote on the true color of a randomly selected jar. The blue jar ($S = B$) contained 11 blue and 9 red balls, the red jar ($S = R$) contained 11 red and 9 blue balls. Jars were chosen with equal probability. Type H voters received a ball with the true color of the jar (the perfectly informative signal), type L voters received a ball that was drawn from the selected jar (the 55% accuracy signal).

Table 3: Experimental Sessions

	N sessions	N matching groups	N subjects
Secrecy	2	5	45
Mild Transparency	2	5	45
Full Transparency	2	4	42

Note: All sessions were run at the DeSciL Lab at the ETH Zurich in May 2013 with students from the ETH and the University of Zurich.

On the first screen, the observer learned that she is an observer, while voters were informed about their type and the color of their ball. Each voter then had to simultane-

⁹In one session we had only 15 subjects and therefore only one matching group.

ously send a message {red, blue, not specified} to the other voter in their group. On the next screen, voters saw the message from their partner and had the opportunity to chat with him for 90 seconds.¹⁰ In the full transparency treatment, the principal could see the voters' messages and follow the chat on her second screen. In the other two treatments the principal could not. On the third screen, the voters could review the communication and then make their final decision by voting for red or blue. Votes were then aggregated to the group decision (red required two votes, blue only one). After the voting stage, the principal received information on her next screen that depended on the treatment. In the secrecy treatment, she could only see the group decision and the true color of the jar. In the mild transparency treatment, she could, in addition, see the individual votes of the voters. In the full transparency treatment, she could also review the whole communication (messages and chat) between the voters. On this screen, the principal had to indicate her belief about the probability that the voters are of type H , by entering this probability in per cent. In the secrecy treatment, in which both voters are indistinguishable to the principal, she had to evaluate one randomly chosen voter from her group, in the other two treatments she had to evaluate both voters in her group. On the final screen of each round, subjects received feedback information regarding the types of the voters, the group decision, the true color of the jar and their pay-offs.

Subjects earned points in each round. The points of a voter in one round was twice the probability that he was a high type, as entered in percentage points by the principal, e.g., if the principal entered 30% the voter's payoff was 60 points. The principal's payoff was 3 points for a correct group decision if the true state of the jar was blue and 1 point for a correct group decision if the true state was red, reflecting the greater importance of a low error rate in the blue state ($x = 0.75$). In addition, the principal earned a number of points between 0 and 100 for accurate evaluation of the voters' types. In each treatment, the evaluation of one of the two voter was randomly selected and a principal j 's earnings were determined by the following quadratic scoring rule:

¹⁰The timeout was not strictly enforced. When the time was up a message appeared on the screen asking them to finish their sentence and proceed. Most subjects did so immediately and the few others were kindly asked to proceed after 120 seconds by an experimenter.

$$\text{Points for accuracy} = \begin{cases} 100 - \frac{1}{100}(100 - Pr_j(t_i = H))^2 & \text{if voter } i \text{ is of type } H \text{ and} \\ 100 - \frac{1}{100}(Pr_j(t_i = H))^2 & \text{if voter } i \text{ is of type } L \end{cases} \quad (6)$$

where $Pr_j(t_i = H)$ denotes the probability that voter i is of type H , as entered by principal j , in per cent. This rule makes it optimal for (risk neutral) subjects to truthfully enter their beliefs (see, e.g., Nyarko and Schotter 2003) and subjects were directly told so in the instructions.

Four rounds were randomly chosen at the end of the session and the points earned in these rounds converted to Swiss Francs at a rate of 1 point= CHF 0.15 (at the time of the experiment CHF 1 was roughly worth USD 1.04). Subjects spend about 2 hours in the lab and earned on average CHF 47 in addition to their show-up fee of CHF 10. Earnings per hour are comparable to an hourly wage for student jobs in Zurich.

4.3. Experimental Results

Aggregate Behavior

Table 4 summarizes the observed error rates and, for comparison, the *ex post* equilibrium predictions, i.e. the predictions after realization of the true state of the world (color of the jar), the types, and the signals. It also contains a column with the hypothetical full information aggregation benchmark case, which assumes perfect information revelation, voting for the state with the higher posterior probability and mixing of committees between voting for (B)lue and (R)ed if the posterior probabilities are equal. This coincides with the equilibrium predictions under secrecy and optimal behavior of a committee that minimizes the total error rate.

Equilibrium predictions are very accurate for the secrecy treatment. In the two transparency treatments, committees make more mistakes than equilibrium predicts in state B (Wald-tests, $p < 0.01$) but do better in state R (Wald-tests, $p < 0.01$ for mild transparency and $p < 0.1$ for full transparency). As predicted, both forms of transparency fare

Table 4: Observed, *Ex post* Equilibrium, and Full Information Aggregation Error Rates by State and in Total

true color	$S = B$			$S = R$			total		
	Observed	Eq.	f.i.a.	Observed	Eq.	f.i.a.	Observed	Eq.	f.i.a.
Secrecy	28.3 (5.5)	27.2	27.2	25.8 (1.7)	24.8	24.8	27 (3.1)	26	26
Mild Transp	15.5 (1.2)	9.5	24	46.1 (3.6)	57.9	25.3	30.8 (1.8)	34	24.7
Full Transp	15.8 (.1)	8.2	21.9	50.7 (6.4)	61.9	26.9	33.3 (3.5)	33.9	24.3

Note: Eq. = likely equilibrium, f.i.a. = full information aggregation. Equilibrium error rates are *ex post* error rates. Standard errors of the observed error rates (in parentheses) are adjusted for clustering in matching groups.

much worse than the information aggregation benchmark in state R , but better in state B . Consequently, the transparent committees performed significantly and substantially worse than secretive committees in state R (Wald-tests, $p < 0.01$) but better in state B (Wald-tests, $p < 0.05$). Even though the total error rate was higher under secrecy than under transparency, principals earned slightly more points in the transparency treatments because state R is more valuable to them ($x = 0.75$). However, this difference was not statistically significant.

The most likely source for the differences between the treatments are (a) differences in information aggregation and (b) coordination behavior of groups consisting only of low types. Information aggregation helps if there is one type H voter and one type L voter in a group, they have conflicting signals and the true state is red. In this scenario all groups in the secrecy treatment aggregated information successfully (see Table 5). While the error rates are substantially and significantly higher (Wald-tests, $p < 0.05$) in the transparency treatments, they are much lower than their predicted value of 100%. In the groups with two low types under transparency, not much coordination on voting for red takes place, resulting in high error rates if the true state is red (however not as high as the predicted 100%) and much fewer errors if it is blue (however more than the predicted 0%). This contrasts with the secrecy treatment where the equilibrium prediction is an error rate of 50% for each state, as these committees are predicted to coordinate on

voting unanimously for red with 50% probability. Indeed, we observe rates closer to 50% under secrecy. As a consequence, error rates are much lower under secrecy if the true state is red and substantially higher if it is blue, which is exactly what principals dislike.

Table 5: Information Aggregation: Error Rates in Groups with Conflicting Signals

true color	$\{H, L\}$ Group	$\{L, L\}$ Group	
	$S = R$	$S = R$	$S = B$
Secrecy	0	54.3 (11.0)	38.7 (15.3)
Mild Transp	22.6 (5.5)	75.7 (4.4)	27.9 (6.8)
Full Transp	44.8 (8.3)	83.3 (10.2)	27.5 (8.9)

Note: Standard errors of the observed error rates (in parentheses) are adjusted for clustering in matching groups.

Surprisingly, not all groups with two signals in the same direction vote for that alternative. As a consequence, the error rates of these groups are not 100% if their signals go in the wrong direction (but range from 82.4% in the mild transparency treatment to 95.6% under secrecy, with 93.3% under full transparency) but also not 0% if they go in the right direction (and range from 4.7% under secrecy to 13.6% under mild transparency, with 7.4% under full transparency, instead).

Individual Behavior

We start our analysis of individual behavior with the deliberation stage. In the straw poll, we see that under secrecy there is almost completely truthful revelation of signals by both high and low types, consistent with equilibrium predictions (Table 6). Under full transparency, high types are also almost always truthful and reveal their signal. However, there are 8.3% low types that lie about their signal and another 10.5% who stay silent, which goes against our predictions. While lying does not make much sense it is also not very costly as L type signals are not very informative. Staying silent might be motivated by the hope to learn more about the true state and then vote accordingly without being punished for it. However, we will see that this does not work out.

Under mild transparency, 19.2% of the high types lie about their signal and another 4.8% stay silent, while the low types are almost always truthful and reveal their signal. The degree of lying from high types is significantly higher than in the other treatments (Wald-tests, $p < 0.01$). It is not high enough though to make the straw poll completely uninformative. Low types could update their beliefs and make fewer mistakes than by sticking to their own signal.

Table 6: (Non-)Truthful messages from different types

	type	lying	silent
Secrecy	<i>H</i> type	0	0
	<i>L</i> type	.7 (.3)	1.8 (1.2)
Mild Transp	<i>H</i> type	19.2 (6.6)	4.8 (1.9)
	<i>L</i> type	3.5 (1.1)	3.3 (1.3)
Full Transp	<i>H</i> type	1.4 (0.9)	0.7 (0.7)
	<i>L</i> type	8.3 (2.1)	10.5 (2.8)

Note: Percentage of non-truthful messages (lying) and “not specified” messages (silent). Standard errors of the observed error rates (in parentheses) are adjusted for clustering in matching groups.

Next, we turn to communication in the chat. Similar to Goeree and Yariv (2011) we observe that communication can roughly be divided into two phases - a first phase, where information is shared, and a second with a discussion on how the committee members should vote. 59.1% of the voters announce a type in the chat (38.2% under secrecy, 49.5% under full transparency and 89% under mild transparency). While announced types are always truthful under secrecy, 51.35% (9.4%) of the claims to be a high type are lies under full (mild) transparency. Claims to be of low type are almost always truthful (96.2% under mild transparency, 98.2 under full transparency and 100% under secrecy). Despite the frequent lies in the full transparency treatment, the announcements are still partially informative and should thus be taken into account in the evaluations by the principals.

Next, we turn to information aggregation again and study how many of the voters who might update their beliefs after receiving a low quality blue signal and seeing the other voter send the message red, indeed vote against their signal in the final vote. The numbers are surprisingly high for the transparency treatments where voting according to signal is predicted (Table 7). Low types always switch under secrecy when it matters most, i.e., when the other group member is a high type, which is facilitated by the truthfulness of announcements of types in the chat. As announcements of types are also truthful most of the time under mild transparency, low types also switch very often under this regime when the other group member is a high type.

Table 7: Information Aggregation: Percentage of Low Types Voting against their blue Signal when other Voter Reports Red

	overall	oth. voter H	oth. voter L
Secrecy	56.2 (6.5)	100	42.6 (11.6)
Mild Transp	45.9 (2.9)	80.0 (5.1)	33.3 (3.7)
Full Transp	41.9 (6.0)	60.0 (5.7)	33.3 (7.2)

Note: Standard errors of the observed error rates (in parentheses) are adjusted for clustering in matching groups.

Next, we study how the principals react. A first look at their evaluations shows that they are well able to distinguish between low types and high types (Table 8). Low types do best under secrecy and high types under full transparency. Average evaluations are too high in all treatments and even significantly positive after wrong decisions (t -tests, $p < 0.01$), suggesting that evaluators do not fully take the relatively low prior into account or that they also care about the pay-offs to the voters. However, the incentives to make a correct group decision under secrecy and a correct individual decision under transparency are about as strong as in our theoretical predictions and always at least 33.5 percentage points.

Under full transparency, evaluators are very well able to distinguish high and low types even if the individual decision is correct. It seems that the principals learn about the types by observing the deliberation. Regressing the evaluations on the pieces of informa-

Table 8: Evaluations

		evaluation		
		avg.	dec. corr.	dec. wr.
Secrecy (group decision)	<i>H</i> type	53.5 (3.4)	53.5 (3.4)	
	<i>L</i> type	37.4 (1.5)	55.0 (3.3)	6.9 (1.9)
	overall	41.6 (1.5)	54.4 (3.2)	6.9 (1.9)
Mild Transp (individual decision)	<i>H</i> type	47.7 (1.2)	48.3 (1.1)	
	<i>L</i> type	31.9 (2.3)	46.4 (2.6)	11.8 (1.6)
	overall	35.7 (1.9)	47.1 (1.9)	11.7 (1.6)
Full Transp (individual decision)	<i>H</i> type	60.8 (3.2)	62.8 (3.8)	
	<i>L</i> type	29.4 (2.9)	40.2 (2.2)	14.6 (4.0)
	overall	37.2 (3.0)	48.2 (1.7)	14.7 (4.0)

Note: Standard errors (in parentheses) are adjusted for clustering in matching groups.

tion that the evaluator sees when evaluating, shows that making the wrong group decision has a big effect on the evaluation under secrecy but no effect under transparency, where she also sees the individual votes (Table 9).

Under transparency, individual mistakes have the biggest influence on the evaluation. Under full transparency, the evaluation is also negatively influenced by switching between the message and the voting stage and if the voter communicates her type to be *L* in the chat. It is positively influenced if the voter communicates her type to be *H*. The principals thus respond to the informational content regarding types in the chat. Staying silent in the straw poll has the same negative effect on evaluations as switching which means that the 10.5% low types who stay silent do not get away with it.

On average, low types who switch between the straw poll and the final vote or stay silent in the straw poll earn 23.6 points less (get an evaluation that is 11.8 percentage points lower) than low types who do not (*t*-test, $p < 0.05$). This result also holds when attention is restricted to the case where the other voter has a different signal.

Table 9: Evaluation Criteria

	M1 (Secrecy)		M2 (Mild Tr.)		M3 (Full Tr.)	
	dependent variable: evaluation ($Pr(t_i = H)$)					
Group decision wrong	-47.6***	(4.2)	-0.9	(3.3)	0.2	(3.2)
Individual decision wrong			-34.8***	(2.1)	-28.7**	(5.4)
Staying silent in straw poll					-12.0**	(2.2)
Switch between m and v					-10.5*	(4.1)
Claimed to be type H in chat					16.4***	(2.1)
Claimed to be type L in chat					-17.4*	(6.5)
constant	54.5***	(3.4)	47.1***	(1.9)	44.3***	(2.3)
N	600		600		560	
N_{clust}	5		5		4	
R^2	0.49		0.31		0.28	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors (in parentheses) are adjusted for clusters in matching groups.

So, principals react to all the information available to them in the right direction, even if the evaluation levels that they choose are too high, on average, to maximise their pay-offs.

5. Conclusion

We show theoretically that transparency can have a strong effect on group decision making. While transparency has a positive effect on the number of correct group decisions in one state of the world it has a negative effect in the other. The question of whether to choose a transparent or secretive committee thus depends on the relative importance the principal attaches to correct decisions in either state of the world. Moreover, our theoretical results suggest that the failure to aggregate information can work through quite different channels, depending on the type of transparency. If the individual voting records

are made public, it works through the incentive not to share information in deliberation. If deliberation is also made public, it works through the incentive to stick to one's announcements even if the posterior probability of the other state is higher after deliberation.

Our experiment is, to our best knowledge, the first experimental study of the effects of career concerns on decision making in groups and our findings mainly confirm our theoretical results. While Goeree and Yariv (2011) find that free-form deliberation eliminates differences in outcomes under different voting rules, our theoretical predictions regarding aggregate outcomes under different levels of transparency largely hold with free-form communication. While these results match the theoretical predictions remarkably well, we do see a considerable level of deviations on the individual level, especially from committee members. As in Goeree and Yariv (2011) and Cai and Wang (2006) players are too truthful in deliberation. Still their behavior corresponds closely enough to our predictions - less truthfulness under mild transparency and less updating and coordination on changing the status quo under both transparency regimes - to lead to drastic differences between secrecy and transparency on the aggregate level.

We thus show that career concerns play an important role and that the way they play out depends crucially on the level of transparency the committee operates under. This suggests that transparency is a highly important element of institutional design and setting the level wrong might have considerable negative consequences for a principal. An interesting extension of our work might be to consider what happens when the level of transparency is set by the committee members rather than by the principal, and whether leaving such a choice to a committee is a good idea.

References

- CAI, H. AND J. T.-Y. WANG (2006): "Overcommunication in strategic information transmission games," *Games and Economic Behavior*, 56, 7–36.
- COUGHLAN, P. (2000): "In Defense of Unanimous Jury Verdicts : Mistrials , Communication , and Strategic Voting," *American Political Science Review*, 94, 375–393.

- FOX, J. AND R. VAN WEELDEN (2012): “Costly Transparency,” *Journal of Public Economics*, 96, 142–150.
- GERSBACH, H. AND V. HAHN (2008): “Should the Individual Voting Records of Central Bankers be Published?” *Social Choice and Welfare*, 30, 655–683.
- GOEREE, J. K. AND L. YARIV (2011): “An Experimental Study of Collective Deliberation,” *Econometrica*, 79, 893–921.
- GUARNASCHELLI, S., R. D. MCKELVEY, AND T. R. PALFREY (2000): “An Experimental Study of Jury Decision Rules,” *American Political Science Review*, 94, 407.
- HOLMSTRÖM, B. (1999): “Managerial Incentive Problems: A Dynamic Perspective,” *Review of Economic Studies*, 66, 169–182.
- IRLENBUSCH, B. AND D. SLIWKA (2006): “Career Concerns in a Simple Experimental Labour Market,” *European Economic Review*, 50, 147–170.
- KOCH, A. K., A. MORGENSTERN, AND P. RAAB (2009): “Career Concerns Incentives: An Experimental Test,” *Journal of Economic Behavior & Organization*, 72, 571–588.
- LEVY, G. (2007): “Decision Making in Committees: Transparency, Reputation, and Voting Rules,” *American Economic Review*, 97, 150–168.
- MEADE, E. E. AND D. STASAVAGE (2008): “Publicity of Debate and the Incentive to Dissent: Evidence from the US Federal Reserve,” *The Economic Journal*, 118, 695–717.
- NYARKO, Y. AND A. SCHOTTER (2003): “An experimental study of belief learning using elicited beliefs,” *Econometrica*, 70, 971–1005.
- PRAT, A. (2005): “The Wrong Kind of Transparency,” *American Economic Review*, 95, 862–877.
- VISSER, B. AND O. H. SWANK (2007): “On Committees of Experts,” *Quarterly Journal of Economics*, 337–372.

——— (2012): “Is Transparency to No Avail?” *Scandinavian Journal of Economics*, 1–34.

Appendix

1. Proofs

Proposition 1. *There always exist truthful equilibria in which each member honestly reveals his signal and type, and the group implements whichever decision has a weighted majority of signals. If the posterior is balanced, the committee randomizes, choosing each state with equal probability. The probability of a wrong decision in each state is given by (5) with thresholds (3) and (4) where $x = 0.5$ in both cases. The posterior beliefs of the principal after correct and incorrect group decisions are given by:*

$$\begin{aligned}\hat{q}_{sec^T}(D \neq S) &= \frac{\sum_{h=0}^n \frac{h}{n} Pr_{sec^T}(D \neq S|h) \binom{n}{h} q^h (1-q)^{n-h}}{Pr_{sec^T}(D \neq S)}, \\ \hat{q}_{sec^T}(D = S) &= \frac{\sum_{h=0}^n \frac{h}{n} (1 - Pr_{sec^T}(D \neq S|h)) \binom{n}{h} q^h (1-q)^{n-h}}{1 - Pr_{sec^T}(D \neq S)}\end{aligned}$$

Proof. If committee members are voting for the posterior most likely state then the principals beliefs must be such that $\hat{q}_{sec^T}(D \neq S) < \hat{q}_{sec^T}(D = S)$. Given these beliefs, and the fact that all players get the same evaluation, the committee will want to choose the state which the group’s collection of signals deem most likely. Deviating from truthful communication or not voting for the most likely state after pooling information would only lower the probability of making the correct group decision, and thus lower expected evaluations. In the case of a balance posterior, the committee would choose each project with probability 0.5. A committee with a balance posterior will be expected to have more low type members than the average committee. As such, if such a committee chose one state more than the other, the principal would incorporate this in his evaluations, and the

resulting payoffs would be lower. A committee with a balanced posterior maximises its payoff by choosing each state with equal probability. \square

Proposition 2. *Both truthful and non-truthful equilibria exist under mild transparency. In any truthful equilibrium the probability of a mistake and the posterior beliefs \hat{q} are the same as in truthful equilibria under secrecy. In any non-truthful equilibrium all committee members vote according to their signal and we have*

$$\begin{aligned}\hat{q}_{mild^{NT}}(v_i = S) &= \frac{q\sigma_H}{q\sigma_H + (1-q)\sigma_L} \\ \hat{q}_{mild^{NT}}(v_i \neq S) &= \frac{q(1-\sigma_H)}{q(1-\sigma_H) + (1-q)(1-\sigma_L)} \\ Pr_{mild^{NT}}(D = R|S = B) &= (1 - q\sigma_H - (1-q)\sigma_L)^n \\ Pr_{mild^{NT}}(D = B|S = R) &= 1 - (q\sigma_H - (1-q)\sigma_L)^n\end{aligned}$$

Proof. The truthful equilibrium has committee members' strategies the same as under secrecy. Supposing all voters are communicating truthfully, we need to check that there is no incentive for a high type voter to mislead the other members. This equilibrium can be sustained with the out-of-equilibrium belief that anyone voting against the majority is a low type. Given this out-of-equilibrium belief the truthful equilibrium is equal to the truthful equilibrium under secrecy.

Now consider the non-truthful equilibrium with beliefs as given in proposition (2). With every member voting their signal, it is clear that $\hat{q}_{mild^T}(D \neq S) < \hat{q}_{mild^T}(D = S)$. In turn, given these beliefs, each member will find it optimal to vote to signal. As communication cannot be believed, there is no profitable deviation in the communication stage. In the final voting stage, each member believes that his own signal is the most likely state of the world, so votes for it. \square

Proposition 3. *Both truthful and non-truthful equilibria exist under full transparency. In both cases all committee members vote according to their signal; the probability of a mistake and the posterior beliefs \hat{q} are the same as in non-truthful equilibria under mild transparency. The truthful equilibrium is sustained by out-of-equilibrium beliefs that any*

members who switches in the final vote from his initial announcement is of type L .¹¹

Proof. The non-truthful equilibrium is as under mild transparency. The principal knows no information is transmitted in the communication stage. It plays no role in the principal's Bayesian updating, so any deviation cannot affect members payoffs.

Consider the truthful equilibrium with beliefs as given in proposition (2). As members vote according to their signals, it must be that $\hat{q}_{full^T}(D \neq S) < \hat{q}_{full^T}(D = S)$. The out-of-equilibrium belief regarding switching makes it therefore optimal to announce the signal truthfully in the straw poll and then not to switch for the same reason that it is optimal to vote according to the individual signal under the non-truthful equilibrium under mild transparency. \square

Proposition 4. *There always exists an x^* such that if $x > x^*$ the principal is better off under transparency than under secrecy.*

Proof. The principal will be better off under transparency than under secrecy when

$$x[Pr_{tran}(D = R|B)] + (1 - x)[Pr_{tran}(D = B|R)] < \\ x[Pr_{sec}(D = R|B)] + (1 - x)[Pr_{sec}(D = B|R)]$$

which can be rearranged to

$$Pr_{tran}(D = R|B) - Pr_{sec}(D = R|B) < \frac{(1 - p)(1 - x)}{px} [Pr_{sec}(D = B|R) - Pr_{tran}(D = B|R)]$$

If x is infinitesimally close to 1, this condition becomes

$$Pr_{tran}(D = R|B) - Pr_{sec}(D = R|B) < -\epsilon \quad (7)$$

Under transparency where players stick to their signals, a mistake will only occur in state B when all players receive incorrect signals in favor of R . This event is a subset of the error events which occur under secrecy in state B , which is why both sides of (7) are negative. \square

¹¹We discuss different assumptions about out-of-equilibrium beliefs and a slightly different truthful equilibrium in Appendix 1.

Proposition 5. (1) Under simple majority rule, the strategies of voters and the beliefs of the principal are the same as under unanimity. Equilibrium error rates will however be different. (2) For any value of x , the principal prefers secrecy to transparency. (3) A principal with $x > x^*$ prefers unanimity with transparency to simple majority with either secrecy or transparency.

Proof. The first part follows from the observation that the equilibria under transparency do not depend at all on the group decision but only on individual decisions and that the committees decide unanimously under secrecy irrespective of the voting rule.

To prove that the principal prefer secrecy to transparency under majority rule we need to show that the following difference is positive

$$\begin{aligned} EU(sec) - EU(tran) &= px(Pr(D \neq B|B, tran) - Pr(D \neq B|B, sec)) \\ &+ (1-p)(1-x)(Pr(D \neq R|R, tran) - Pr(D \neq R|R, sec)) \end{aligned} \quad (8)$$

In a committee of size n , we know that under transparency the group will make the wrong decision whenever a majority of players have the wrong signal.

$$Pr(D \neq S|tran) = \sum_{k=\frac{n+1}{2}}^n \binom{n}{k} (q(1 - \sigma_H) + (1 - q)(1 - \sigma_L))^k (q\sigma_H + (1 - q)\sigma_L)^{n-k} \quad (9)$$

The probability of making a mistake under secrecy is slightly different from that of transparency. Under secrecy the group may actually follow a minority of signals if there are enough H types in the minority. This leads to less mistakes when the minority is correct, but more when it is wrong. The difference between secrecy and transparency in terms of utility corresponds exactly to these cases where a group under secrecy goes against the majority. We get

$$\begin{aligned} EU(sec) - EU(tran) &= \sum_{i=1}^{\frac{n-1}{2}} x(Pr(D = B|S = B, \#_{s_B} = i) - Pr(D = R|S = B, \#_{s_B} = n - i)) \\ &+ (1 - x)(Pr(D = R|S = R, \#_{s_B} = n - i) - Pr(D = B|S = R, \#_{s_B} = i)) \end{aligned}$$

where for a particular number of signals $i < \frac{n+1}{2}$ in favour of state B we have

$$Pr(D = B | S = B, \#s_B = i) = \sum_{j=1}^{2i-1} \sum_{k=0}^{\frac{j-1}{2}} \binom{n}{j} q^j (1-q)^{n-j} \binom{j}{j-k} \sigma_H^{j-k} (1-\sigma_H)^k \binom{n-j}{i-j+k} \sigma_L^{i-j+k} (1-\sigma_L)^{n-i-k}$$

$$Pr(D = B | S = R, \#s_A = i) = \sum_{j=1}^{2i-1} \sum_{k=0}^{\frac{j-1}{2}} \binom{n}{j} q^j (1-q)^{n-j} \binom{j}{j-k} \sigma_H^k (1-\sigma_H)^{j-k} \binom{n-j}{i-j+k} \sigma_L^{n-i-k} (1-\sigma_L)^{i-j+k}$$

The probabilities above are made up of many (i, j, k) combinations, where i is the realised number of signals in favour of B , j is the realised number of H types, and k is the realised number of R signals from H types. To prove the proposition it suffices to consider a subset of events. Take the event $\#s(A, R) = (i, n-i)$ so there are i signals in favour of state A and $n-i$ signals in favour of R , where $i < \frac{n+1}{2}$. The mirror of this event is $\#s(A, R) = (n-i, i)$. There are two different cases to consider for a particular (i, j, k) combination.

1. If $(\frac{\sigma_H}{1-\sigma_H})^{2k-j} (\frac{\sigma_L}{1-\sigma_L})^{n-2i-2k+j} > 1$ then in both $\#s(A, R) = (i, n-i)$ and $\#s(A, R) = (n-i, i)$ we have $EU(sec) = EU(tran)$. That is, the minority do not have strong enough signals to turn the decision in their favor.
2. If $1 > (\frac{\sigma_H}{1-\sigma_H})^{2k-j} (\frac{\sigma_L}{1-\sigma_L})^{n-2i-2k+j}$ then

$$\begin{aligned} EU(sec) - EU(tran) = & \binom{n}{j} q^j (1-q)^{n-j} \binom{j}{j-k} \binom{n-j}{i-j+k} \\ & x(\sigma_H^{j-k} (1-\sigma_H)^k \sigma_L^{i-j+k} (1-\sigma_L)^{n-i-k} - \sigma_H^k (1-\sigma_H)^{j-k} \sigma_L^{n-i-k} (1-\sigma_L)^{i-j+k}) + \\ & (1-x)(\sigma_H^{j-k} (1-\sigma_H)^k \sigma_L^{i-j+k} (1-\sigma_L)^{n-i-k} - \sigma_H^k (1-\sigma_H)^{j-k} \sigma_L^{n-i-k} (1-\sigma_L)^{i-j+k}) \end{aligned}$$

This is always positive.

Part (3) follows directly from (1) and (2). As committees under secrecy decide unanimously the error rates are the same under both voting rules and the (3) follows from the fact that the error rates under transparency and unanimity are preferred to the error rates under secrecy by assumption. \square

2. Out-of-equilibrium Beliefs

In proposition 3, we assume the out-of-equilibrium belief that a principal believes a committee member that switches between the deliberation message m and the final vote v is of type L . This belief is intuitive as under many straw poll results type H voters who find themselves in the minority would still hold the posterior belief that the true state is indicated by their signal, while type L voters' posterior beliefs can be changed by a majority of opposing messages more easily. So, while high types have no reason to switch after many straw poll results, low types would have an incentive to switch, e.g. if they mistakenly decide to help the principal instead of improving their reputation. This argument could be made in a more rigorous and formal way by assuming an ϵ probability that committee members care about the principal's pay-off rather than about their reputation. Type H committee members of this type would still not want to switch but Type L committee members would want to switch. This would make switching a non-zero probability event in equilibrium where switchers are then low types with certainty.

However, if the number of majority messages in deliberation is high enough also high types who find themselves in the minority will change their belief about which state is more likely. In this case a different equilibrium is equally plausible. In this equilibrium, neither high nor low types in the minority can change their reputation after a straw poll with sufficiently many majority messages, and these members switch in the final vote to help the principal. The results of Proposition 3 would change as follows.

Let θ denote the number of majority signals in the straw poll and θ^* the threshold of majority signals above which high and low types who supported B in deliberation and find themselves in the minority would want to switch:

$$\theta^* = \frac{n-1}{2} + \frac{\log\left(\frac{\sigma_H}{(1-\sigma_H)}\right)}{\log\left(\frac{q(\sigma_H-\sigma_L)+\sigma_L}{1-q(\sigma_H-\sigma_L)-\sigma_L}\right)} \quad (10)$$

Proposition 6. *Both truthful and non-truthful equilibria exist under full transparency. In any non-truthful equilibrium the probability of a mistake and the posterior beliefs \hat{q} are the same as those in the non-truthful equilibria under mild transparency. In any truthful equilibrium (i) if $\theta < \theta^*$ then the probability of a mistake and posterior beliefs will be as in the non-truthful equilibrium, (ii) if $\theta > \theta^*$ then posterior beliefs are given by*

$$\begin{aligned} \hat{q}_{fullNT}(i \in maj, m_i = v_i = S) &= \binom{n}{\theta} \frac{q\sigma_H(q\sigma_H + (1-q)\sigma_L)^\theta (1-q\sigma_H - (1-q)\sigma_L)^{n-\theta}}{q\sigma_H + (1-q)\sigma_L} \\ \hat{q}_{fullNT}(i \in maj, m_i = v_i \neq S) &= \binom{n}{\theta} \frac{q(1-\sigma_H)(q\sigma_H + (1-q)\sigma_L)^{n-\theta} (1-q\sigma_H - (1-q)\sigma_L)^\theta}{q(1-\sigma_H) + (1-q)(1-\sigma_L)} \\ \hat{q}_{fullNT}(i \in min, m_i \neq S) &= \binom{n}{\theta} \frac{q(1-\sigma_H)(q\sigma_H + (1-q)\sigma_L)^\theta (1-q\sigma_H - (1-q)\sigma_L)^{n-\theta}}{q(1-\sigma_H) + (1-q)(1-\sigma_L)} \\ \hat{q}_{fullNT}(i \in min, m_i = S) &= \binom{n}{\theta} \frac{q\sigma_H(q\sigma_H + (1-q)\sigma_L)^{n-\theta} (1-q\sigma_H - (1-q)\sigma_L)^\theta}{q\sigma_H + (1-q)\sigma_L} \end{aligned}$$

This equilibrium is sustained by out-of-equilibrium beliefs that any member who switches from his initial announcement in the final vote is of type L if $\theta < \theta^*$. The resulting error probabilities are:

$$\begin{aligned} Pr(D = R | S = B) &= \sum_{\theta=\theta^*}^n \binom{n}{\theta} (1-q\sigma_H - (1-q)\sigma_L)^\theta (q\sigma_H + (1-q)\sigma_L)^{n-\theta} \\ Pr(D = B | S = R) &= 1 - \sum_{\theta=\theta^*}^n \binom{n}{\theta} (q\sigma_H + (1-q)\sigma_L)^\theta (1-q\sigma_H - (1-q)\sigma_L)^{n-\theta} \end{aligned}$$

Mistakes will happen in state B when $\theta \geq \theta^*$ voters receive the wrong signal. Mistakes will happen in state R whenever less than θ^* voters get R signals.

Note that Proposition 4 still holds, just for a different threshold for x in this case. Note further that for $\sigma_H = 1$, the parameter choice in the experiment, $\theta < \theta^*$ always holds.

3. Instructions to Experiment

Instructions for the full transparency treatment (translation; original in German). Instructions for the other treatments were very similar and are therefore omitted here.

Overview

Welcome to this experiment. We kindly ask you not to communicate with other participants during the experiment and to switch off your phones and other mobile devices.

At the end of the experiment you will be paid out in cash for your participation in today's session. The amount of your pay-off depends in parts on your decisions, on the decisions of other participants and on chance. For this reason it is important that you read the instructions carefully and understand them before the start of the experiment.

In this experiment all interactions between participants are via the computers that you are sitting in front of. You will interact anonymously and your decisions will only be stored together with your random ID number. Neither your name, nor names of other participants will be made public, not today and not in future written analyses.

Today's session consists of several rounds. **At the end, 4 rounds will be randomly selected and paid out. The rounds that are not chosen will not be paid out.** Your pay-off results from the points that you earn in the selected rounds, converted to Swiss Francs, plus your show-up fee of CHF 10. The conversion of points to Swiss Francs happens as follows: Every point is worth 15 cents, which means that

20 points = CHF 3.00.

Every participant will be paid out in private at the payment counter, so that no other participant can see how much you have earned.

Experiment

This experiment consists of 20 procedurally identical rounds. In each round a group decision has to be made, that can be correct or wrong.

Two members in each group of three make the group decision (henceforth we will call them the voters). There are well and less well informed voters and the task of the third group member is to observe the decision process of the other two members and then to indicate the probability with which he thinks that the other group members are well or less well informed (henceforth, we will call this member the observer).

The higher the evaluation of the observer with respect to the level of information of a voter is, the higher is the pay-off to that voter in the round. The more accurate the evaluation of the observer with respect to the level of information of the voters is, the higher is his or her pay-off in the round. In addition, the observer receives a pay-off for correct group decisions.

The Group

In the **first** round you will be assigned a meta-group of 9 members. In the beginning of **every** round you will be randomly assigned to a **new group** which consists of randomly selected members of your meta-group. Every group has three members: 2 voters and 1 observer.

Whether you will be assigned the role of a voter or an observer, is randomly determined each round. The voters receive, again randomly, the labels “voter 1” and “voter 2”.

All interactions in a round take place within your group of three.

The Voters

There are two types of voters, well informed (type G) and (less well) informed (type I) voters. Of which type the group members are, is again determined randomly. With probability $\frac{1}{4}$ (or 25%) a voter receives good information which means he is of type G; with probability $\frac{3}{4}$ (or 75%) he receives less good information which means he is of type I.

Because the assignment of types to the voters is independent of the assignment to other voters, there can be two voters of type G, two voters of type I, or one of each type in a group.

The voters learn their type on the first screen of a round but not the type of the other voter in their group. The observer learns that he is an observer on the first screen but not the types of the voters in his group.

Later, after observing the behavior of the voters, it will be the task of the observer to estimate the probabilities that voter 1 and voter 2 are of type G.

The Jar

There are two jars: one red jar and one blue jar. The red jar contains 11 red and 9 blue balls, the blue jar 11 blue and 9 red balls. Each round one jar will be randomly selected.

The task of the voters is to vote on the color of the jar. Each jar has an equal probability of being selected, that is it will be selected with 50% probability.

The Ball

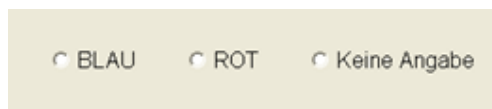
The well informed voters (type G) receive a ball with the actual color of the jar, that is they are directly informed about the color of the jar.

The informed voters (type I) receive a randomly drawn ball from the selected jar. They are not told the color of the jar. If there are two type I voters in a group, each of them receives a ball from the jar. Every ball in the jar has the same selection probability for the type I voters, that is for each voter of type I a ball is drawn from a jar containing 20 balls (11 with the color of the jar, 9 with the other color).

The voters learn the color of their ball on the first screen. Every voter only sees the color of his ball, not the color of the other voter's ball.

Communication

After learning their type and the color of their ball, the voters can communicate the color of their ball to the other voter in their group. They can also communicate the color that their ball did not have or stay silent. The communication is made through the following entry mask.



☐ BLAU ☐ ROT ☐ Keine Angabe

On the following screen the voters learn the message of the other voter in their group and have the option to chat with him. The chat happens via the following entry mask.

Sie können nun mit dem anderen Wähler in ihrer Gruppe chatten. Im Chat bekommen Sie die ID "Wähler 1" und der andere Wähler die ID "Wähler 2". Jede Texteingabe muss mit der Enter-Taste bestätigt werden.

Wähler 2: Hallo
Wähler 1: Chat Test

You can enter arbitrary text messages into the blue entry field. Pay attention to confirm every entry by pressing the enter button to make it visible for the other voter. It will then appear in the grey field above.

The observer cannot participate in the communication but sees the messages of the two voters regarding the color of their ball as well as the chat.

Group Decision

After the communication stage the voter make their decision in a group vote.

So, if you are a voter, you have to vote either for blue or for red.

Once both voters have made their decision, the votes for blue and red are added up and the group decision results from the following rule:

- **If the color RED receives 2 votes, the group decision is RED**
- **If the color BLUE receives 1 or 2 votes, the group decision is BLUE**

That is for a group decision for blue only one vote is necessary while a group decision for red requires two votes.

Evaluation of the Observer

After the voters have cast their vote and the group decision is determined, the evaluator learns the group decision as well as the decisions of the individual voters in his group.

Moreover, he learns the true color of the jar, that is, whether the group decision and the individual decisions were correct or wrong.

On the same screen the observer can review the entire communication between the voters in his group once again.

If you are an observer, you now have to enter for each of the two voters the probability with which you believe that this voter is of type G.

To do so you enter a number between 0 and 100 which expresses your evaluation in percentage points. The entry mask looks as follows.

Geben Sie die Wahrscheinlichkeit in Prozent an, mit der Sie glauben, dass **Wähler 1** ein Typ X ist (eine Zahl zwischen 0 und 100):

The complete screen of the observer looks as follows (example screen).

Die Dose hat die Farbe **BLAU**

Die Wähler haben wie folgt abgestimmt

Wähler 1 hat für folgende Farbe abgestimmt: ROT	Wähler 2 hat für folgende Farbe abgestimmt: ROT
---	---

Die Gruppenentscheidung ist somit: **ROT**

Hier sehen Sie die Kommunikation zwischen den Wählern noch einmal

Wähler 1 gab nicht an, welche Farbe seine Kugel hat. Wähler 1: Hallo Wähler 1: Text Chat	Die Kugel, die Wähler 2 angab, ist: ROT
--	--

Geben Sie nun Ihre Einschätzung zu Wähler 1 und Wähler 2 an.

Geben Sie die Wahrscheinlichkeit in Prozent an, mit der Sie glauben, dass Wähler 1 ein Typ G ist (eine Zahl zwischen 0 und 100): <input type="text"/>	Geben Sie die Wahrscheinlichkeit in Prozent an, mit der Sie glauben, dass Wähler 2 ein Typ G ist (eine Zahl zwischen 0 und 100): <input type="text"/>
---	---

Weiter

Pay-off in each Round

If you are a **voter** your pay-off is determined by the evaluation of the observer. If the observer believes that you are of type G with **X%** probability, you receive a pay-off of **2*X** points in

this round. This means that your pay-off directly depends on the probability with which the observer believes you are a **well-informed voter (type G)**.

If the observer has entered the probability 25%, for example, your pay-off is 50 points, if he has entered 50%, it is 100 points.

If you are an **observer** you receive a pay-off for correct group decisions and a pay-off for the accuracy of your evaluations of the types of the voters.

- If the group decision is RED and the jar is indeed RED, you as an observer receive **1 point**.
- If the group decision is BLUE and the jar is indeed BLUE, you as an observer receive **3 points**.
- If the group decision is wrong, you receive **0 points**, independently of the true color of the jar.

For your evaluation regarding the types of the voters you receive a pay-off between 0 and 100 points. It will be randomly determined whether you will be paid out for the evaluation of voter 1 or voter 2.

If you have evaluated both voters correctly with certainty (that is with 0 or 100%) (if you entered the probability 0 for both voters, for example, and both are indeed not of type G but of type I), you receive 100 points. If you are completely wrong (if both are of type G in the example) you receive 0 points.

The formula that determines your pay-off is a little complicated.

Put simple the formula assures **that it is best for you (gives you the highest expected pay-off) if you truthfully indicate the probability with which you believe that a voter is indeed of type G**. Every other evaluation lowers your expected pay-off.

If you believe, for example, that voter 1 in your group is of type G with 30% probability and voter 2 with 60% probability, it is best for you to enter exactly these values.

In case you want to in more detail how your payoff is determined: for the evaluation of the randomly selected voter you receive:

$$100 - \frac{1}{100} (100 - \text{prob}(\text{voter is of type G}))^2, \text{ if this voter is of type G and}$$
$$100 - \frac{1}{100} (\text{prob}(\text{voter is of type G}))^2, \text{ if this voter is of type I,}$$

where $(\text{prob}(\text{voter is of type G}))$ is your indication of probability in percentage points that that voter is of type G. The resulting number is rounded up to a whole number and gives, together with your pay-off in case of a correct group decision, your pay-off in the round.

Remember: At the end of the experiment 4 rounds are randomly selected, the point incomes converted to Swiss Francs and paid out in private. The not selected rounds will not be paid out.

Questions?

Take your time to read the instructions carefully. If you have any questions, raise your hand. An experimental administrator will then come to your seat.