**Extended Abstract**

We consider the classical problem of information transmission between a sender with private payoff-relevant information and a receiver who takes actions which also affect the sender's payoff. Following the Bayesian persuasion literature pioneered by Rayo and Segal (2010) and Kamenica and Gentzkow (2011), we suppose that the sender has commitment power over the information she reveals to the receiver. Without setting any restrictions on possible persuasion strategies, we search for conditions under which full disclosure is optimal. Differently from other complicated schemes, just disclosing the truth seems to be a realistic goal in many scenarios – e.g., with transparency policies in organizations.

In our model, the state space can be a continuum, therefore the concavification approach of Kamenica and Gentzkow (2011) is not operational. Moreover, differently from several recent papers (e.g., Dworczak and Martini (2019), Dizdar and Kováč (2020), Gentzkow and Kamenica (2016), Kolotilin (2018), Kolotilin et al. (2021), Arieli et al. (2020)), we do not assume that the sender's payoff is a function of the expected state (or any moments of the posterior distribution). Despite this, we obtain a sufficient condition for the optimality of full disclosure that speaks directly to the underlying incentives of the parties, as opposed to the indirect utility function of the sender. This makes our condition easily interpretable and verifiable. In particular, it can be interpreted as a requirement of minimal alignment of incentives between the sender and the receiver.

To see why the effect of full disclosure may be non-trivial, consider a simple principal-agent setup, as an example. The agent generates an output which he and the principal split in a fixed proportion. The output is increasing in the agent's effort, and the agent bears the cost of effort. The underlying state of nature determines the productivity of effort, with a higher state resulting in higher productivity. The principal knows the state, while the agent does not. At first sight, the principal would always want to commit to revealing the state to the agent, as both parties seem to benefit from effort more when the state is higher. Here is a simple argument why it does not have to be the case. Suppose the agent is sufficiently risk averse. Then, good news about the productivity may actually depress effort. This is because a higher productivity implies that the agent reaches a higher income, hence a lower marginal utility, at lower levels of effort. If the principal is risk neutral, then disclosure discourages the agent precisely when the principal benefits more from effort (and incentivizes the agent when the principal gains less from effort). In such a case, transparency is unlikely to be optimal. Note also that, even when this "income effect" does not prevail in the agent's incentives, full disclosure may not be optimal. Even if the agent increases effort under the good news and reduce it under the bad news about the state, as the principal wants, the increase may be smaller than the decrease to the point that the overall effect on the principal's utility is negative.

To see how we tackle these difficulties, stick to the principal-agent setup and consider the following example. Fix a message that pools two ex-ante equally likely states. The principal contemplates splitting this message into two messages that reveal the state. Then, given the optimal effort under the pooling message, the agent will discover that her marginal utility of effort is positive when one state is revealed, negative when the other state is revealed, and the two values have the same magnitude, just opposite signs. Thus, the agent will decrease effort under the first state and increase it under the second state. Two forces determine whether the principal gains from the split or not: the changes in the agent's effort and the changes in the principal's utility per unit of effort. Under each state, the agent modifies his effort until its marginal utility returns to zero. Then what matters is how much the principal's utility increases (or decreases) per unitary increase (or decrease) of the agent's marginal utility. In particular, the principal benefits from the split if this measure of her marginal utility is larger when the agent is happy to increase effort with respect to when the agent prefers to reduce effort. In this sense, ours is a condition of minimal alignment of interest between the two parties.

Our main result extends this argument to all possible messages in a general sender-receiver framework. Specifically, in Section 3, we show that any message with a non-singleton support can be split so as to improve the sender's welfare if a decrease of the receiver's marginal utility by one unit has a larger benefit for the sender when it also benefits the receiver, compared to when it harms him. This condition, hence, ensures the optimality of full disclosure.

A comparison of our condition with two conditions obtained in the literature for the special case when the receiver's action is linear in expected state suggests that ours is rather weak. It coincides with the necessary and sufficient condition when the sender's utility depends only on the receiver's action and is weaker that the sufficient condition provided in Kolotilin et al. (2022) when the sender's utility is allowed to depend on the state as well as the action.

Under the assumption of continuity of the state space and some additional regularity requirements, we also derive a sufficient condition in terms of just derivatives of the parties' utility functions ("derivatives condition"). This condition is slightly stronger than our general condition but may be easier to check in some economic applications.

Finally, we also derive a sufficient condition for suboptimality of full disclosure. Although there remains "a gap" between this condition and our optimality condition (one is not a negation of the other), it still helps to establish when full disclosure is definitely not optimal, as we show later in an example.

We then turn to a principal-agent setting outlined in the third paragraph of this section and discuss several examples demonstrating that our optimality conditions are easy to check and often satisfied in economic applications. The first example (section 5.1) sheds light on the role of risk aversion for the optimality/suboptimality of full disclosure. We assume that both parties exhibit CRRA. Full disclosure turns out to be optimal, when the agent is more risk averse than the principal (a typical textbook situation) but not too risk averse (with the coefficient of relative risk aversion below one). In this case, the preferences of the parties are sufficiently aligned. For both parties, state and effort remain complements. Disclosing the states raises the effort in expectation, and the principal benefits more from effort in higher states. Instead, when the agent becomes too risk averse (while the principal remains moderately risk averse), full disclosure ceases to be optimal. As we have discussed in the beginning in this section, under high agent's risk aversion, good news about productivity depress effort, that is, effort and state become substitutes for the agent while remaining complements for the principal.

Another interesting, though less realistic, case discussed in Section 5.1 is when the agent is sufficiently risk averse, and the principal is at least as risk averse as the agent. In that case, the average effort falls but the principal nevertheless gains from transparency. This happens because for the principal effort and state are even more substitutes than for the agent. Bad news about productivity encourages effort, and the principal benefits even more from effort in lower states than the agent does.

In the second example (section 5.2) we simplify preferences by assuming risk neutrality for both parties and focus instead on the properties of the production function that ensure the optimality of full disclosure. By applying the "derivatives version" of our sufficient condition we show that full disclosure is optimal under some commonly used functional forms for output.

Two recent papers, Kolotilin (2018) and Kolotilin et al. (2022) have provided a general necessary and sufficient condition for the optimality of full disclosure in persuasion. This condition is expressed in terms of the sender's indirect utility function (given the receiver's optimal choice of action) and essentially requires that, for any pair of states, the sender prefers their revelation to any pooling of them.

Except that it reduces the problem to checking only messages with a binary support, it remains a "high level" condition that is hard to operate, because it requires veryfing it for all possible pairs of states and posteriors (which also requires calculating the receiver's optimal response to any posterior). In the special case when the sender's payoff can be represented as a function of only the posterior mean, the necessary and sufficient condition for the optimality of full disclosure boils down to the requirement of the convexity of this function (see, e.g., Kolotilin (2018) and Kolotilin et al. (2022)).

In our model, except for some regularity assumptions and the requirement that the receiver's utility is concave in action and delivers an interior solution, we impose no restrictions on how the state and the action affect utilities, and we offer a condition in terms of primitives of the model, which relates the incentives of the sender and the incentives of the receiver. All in all, compared to the general condition in Kolotilin et al. (2022), we lose "necessity" but gain "operability".

Using a concept analogous to the concept of "virtual value" in the mechanism design literature, Mensch (2021) offers conditions for full disclosure jointly on the receiver's utility function and on a transformation of the sender's utility function that takes into account the incentive compatibility constraint of the receiver ("virtual utility"). His focus is on the importance of complementarities between states and actions, and whether these complementarities "point in the same direction" for the sender and the receiver. Our and Mensch's papers can be viewed as complementary to each other. While Mensch's condition for full disclosure (Theorem 5) is intuitive, it is rather abstract and not straightforward to apply, as it requires a derivation of the "virtual utility". Instead, our conditions are directly on primitives of the model, that is, the shapes of the parties' utility functions.

**References:**

Arieli I., Y. Babichenko, R. Smorodinsky, and T. Yamashita, "Optimal persuasion via bi-pooling," working paper, available at https://ssrn.com/abstract=3511516.

Dizdar D., and E. Kováč, 2020, "A simple proof of strong duality in the linear persuasion problem," Games and Economic Behavior, 122: 407-12.

Dworczak, P., and A. Kolotilin, 2019, "The persuasion duality," working paper, available at https://ssrn.com/abstract=3474376.

Dworczak P, and G. Martini, 2019, "The simple economics of optimal persuasion," Journal of Political Economy, 127(5): 1993-2048.

Gentzkow M, and E. Kamenica, 2016, "A Rothschild-Stiglitz approach to Bayesian persuasion," American Economic Review, 106(5): 597-601.

Kamenica E., and M. Gentzkow, 2011, "Bayesian persuasion," American Economic Review, 101(6): 2590-615.

Kolotilin, A., 2018, "Optimal information disclosure: A linear programming approach," Theoretical Economics, 13(2) :607-35.

Kolotilin, A., R. Corrao, and A. Wolitzky, 2022, "Persuasion with non-linear preferences," UNSW Economics Working Paper 2022-03, available at https://ssrn.com/abstract=4204824.

Kolotilin, A., T. Mylovanov, and A. Zapechelnyuk, 2021, "Censorship as optimal persuasion," working paper.

Mensch J., 2021, "Monotone persuasion," Games and Economic Behavior, 130: 521-42.

Rayo L, and I. Segal, 2010, "Optimal information disclosure," Journal of Political Economy, 118(5): 949-87.