Persuasion with Multiple Actions*

Davit Khantadze † Ilan Kremer‡ Andrzej Skrzypacz§

Abstract

We consider a Bayesian persuasion model in which multiple receivers take one action each. We compare simultaneous procedures to sequential ones. In a simultaneous procedure, all the receivers act simultaneously following the realization of a single public signal. In a sequential procedure, receivers receive information and take actions sequentially. We characterize the conditions under which the optimal sequential procedure leads to a higher payoff and characterize the optimal ordering of actions.

JEL Codes: D21

1 Introduction

The literature on Bayesian persuasion that began with Kamenica and Gentzkow (2011) characterizes how a sender affects a receiver's action by designing a public signal. With a verifiable public signal, the sender can commit to revealing the signal regardless of its realization. While such commitment is beneficial for the sender, it poses a challenge when the sender wishes to affect the decisions of multiple agents. A public signal that leads to the optimal action (from the sender's perspective) of one agent would typically lead to a suboptimal action by another agent. A natural strategy would then be to provide information gradually and let agents act sequentially. This would enable the sender to modify the information the subsequent agent observes before taking his action.[1]

*We thank seminar participants, our colleagues, and students for their comments and suggestions. This research was made possible by the generous support of NSF Grant # 2116250 and BSF Grant # 2021642. Edited by Emir Kamenica.

†Independent Researcher

‡The Hebrew University and University of Warwick

§Stanford Graduate School of Business, Stanford University

[1]See the discussion in Kamenica and Gentzkow (2011) and Kamenica (2019) about the many ways the sender may commit to the signal. In most applications discussed in those papers, public information disclosure helps with commitment. As we discuss later, our results can be interpreted as how sequential persuasion can mitigate the costs of publicity.

An example of sequential persuasion is a multinational company that can come under investigation for some conduct (for example, alleged market power) in multiple jurisdictions. The conduct of that company can be different in different countries but is likely correlated. The regulators and judges in every country may follow different laws when deciding the cases. When the regulators move sequentially, information revealed by the company in one case becomes available for the subsequent regulators. It can become either publicly available (in the form of the report of the early-moving regulators) or the existence of a source can become available, and the subsequent regulators can request the same data as produced in the previous cases.[2]

Another example is when a prosecutor tries to obtain multiple convictions for the same crime. A known case is the trial for the murder of Mrs. Stout. In the initial trial, the prosecution asserted that Mr. Stumpf had played the primary role in Mrs. Stout's murder, resulting in a death sentence for Mr. Stumpf. Several months later, during Wesley's trial, the same prosecutor from Mr. Stumpf's case, before the same judge who had presided over Mr. Stumpf's trial, presented evidence suggesting that it was Mr. Wesley, not Mr. Stumpf, who had shot and killed Mrs. Stout. After several appeals that reached the US Supreme Court, it was finally decided that both death sentences were upheld. As noted in a 2017 article in The New Yorker, '...a defendant can be convicted of murder without being the killer. But, if the prosecutor says that a defendant pulled the trigger, it's easier to ask a judge or a jury for a death sentence' [3]. Anthony-North (2022) finds that this is not an isolated case, there are at least 29 similar death sentences. She writes: "This case study illustrates a practice among prosecutors of a single sovereign whereby they pursue incompatible theories of a case against two or more defendants for criminal behavior for which, factually, only one defendant can be culpable."

Apart from the examples mentioned above, there are other situations where senders may benefit from sequential persuasion. Such examples include public disclosure of information by politicians who try to gain support for a sequence of bills, committee heads who try to influence votes on multiple issues, and managers trying to convince their budget committees to provide funding for multiple projects.

[2]For example, in June 2021, Google settled with the French competition authority regarding the display advertising market (https://www.theguardian.com/technology/2021/jun/07/france-fines-google-for-abusing-online-advertising-dominance). The UK's Competition and Markets Authority has opened an investigation of Google's conduct in that market in May 2022, with " Initial investigation: including information gathering, analysis and review of information gathered" taking place from May 22 to February 2023 (https://www.gov.uk/cma-cases/investigation-into-suspected-anti-competitive-conduct-by-google-in-ad-tech). The US DOJ brought charges against Google's conduct in this market in January 2023, while the European Commission in June 2023 (https://www.nytimes.com/2023/06/14/technology/google-antitrust-european-union.html?smid=url-share)

[3]https://www.newyorker.com/magazine/2017/11/13/two-murder-convictions-for-one-fatal-shot

To illustrate our model, consider the classic example of Bayesian persuasion from the introduction of Kamenica and Gentzkow (2011). They consider a prosecutor (sender) who wishes to maximize the probability of conviction by the judge (receiver). The judge finds a defendant guilty if and only if his belief that the defendant committed a crime is above a certain threshold. Consider instead two judges who decide whether to convict two defendants, A and B. The two defendants are accused of two different crimes, and the prior of them committing these crimes is negatively correlated. The prosecutor who wishes to maximize the expected number of convictions needs to account for the fact that a signal regarding A also affects the second judge who decides on B. If the two judges decide simultaneously based on a single signal, the prosecutor faces a potential tradeoff. Given the negative correlation, increasing the probability of the conviction of A may decrease the probability of the conviction of B and vice-versa.

A sequential procedure with two steps could help. In the first step, the prosecutor provides one public signal, and the first judge decides on A. In the second step, the prosecutor reveals the realization of an additional signal, and the second judge decides on B. The main questions we address in this paper:

- Do sequential procedures dominate simultaneous ones?
- Assuming that the answer to the first question is yes, under what conditions is the sequential procedure strictly better than the simultaneous procedure?
- Given a sequential procedure, what order is optimal for the sender?
- To what extent can the sender benefit from flexibility where the order of the sequence can be based on the realized signals?

In our baseline model, there is an arbitrary number of receivers that we denote by n.[4] There are n binary states with an arbitrary correlation structure. Receiver i takes action a_i, which is some function of his belief regarding state i. The sender maximizes the sum of payoffs from the receivers' actions where the payoff following a single receiver's action is some arbitrary function of his action (that can be state-dependent).

We first note that the sender weakly prefers a sequential procedure to a simultaneous one. This is because the sender can in each round stay quiet or choose uninformative signals. Hence, we examine when a sequential procedure strictly outperforms a simultaneous one.

[4] Another interpretation of our model is that of a single receiver who takes multiple actions. This interpretation assumes that the receiver does not obtain additional information following the actions he takes. The only source of information is what is provided by the sender. This assumption is realistic, for example, in the case of a hiring committee making offers to multiple candidates and the quality of the candidates revealed after they start working.

We also show that there is an upper bound for the payoff that the receiver can achieve that applies to any procedure, sequential or simultaneous. The upper bound is constructed by considering the minimal information leakage. This payoff would be obtained if the signals were sent via private communication. We refer to that upper bound as the "first-best" payoff.

In Section 3, we consider a sequential procedure where the sender predetermines the order in which receivers move. That is, the order cannot be modified based on the realization of signals. Such a restriction is relevant in many cases, such as the above-mentioned legal example where discovery happens only after a motion is filed. We show that the optimal sequential mechanism strictly improves the sender's payoff if and only if the simultaneous procedure fails to achieve the first-best payoff. It is important to note that when the order of actions is predetermined, the sequential mechanism may fail to achieve the first-best outcome. Nevertheless, even when it fails to achieve the upper bound, it generates strictly higher payoffs than the simultaneous procedure. Intuitively, if simultaneous persuasion performs strictly worse than private persuasion, then one can find a receiver who is not given optimal information. If he is provided with "too much" information, one can benefit from making him make the decision first and provide him with less information. If he is provided with "too little" information, one can benefit from making him make the decision last and provide him with additional information. In the Appendix, we present a complete characterization of a simple set of examples that allows us to characterize when the simultaneous procedure fails to achieve the first-best payoff and what the optimal order should be.

In Section 4, we deviate from the assumption that the order is predetermined. As we later discuss, one interpretation of this setup revolves around the concept that the sender isn't the originator of new information; rather, it emanates from an external, exogenous source. The sender's role primarily involves timing the actions of the receivers. At each point in time, the sender determines whether to prompt a receiver to take action and, if so, whom to prompt.[5] We demonstrate that if the sender can select the sequence of receivers in response to the signal's realization, the sender can always attain the optimal, first-best payoff.

As mentioned before, our baseline model assumes states to be binary, while actions may encompass more than two values. The sender's payoff is assumed to be additive. In Section 6, we investigate a scenario in which actions are binary (e.g., accept/reject) but relax other assumptions. This exploration enables us to delineate the optimal sequence and consider situations where the sender endeavors to persuade multiple voters, necessitating the attainment of a specific majority. Additionally, we assess the extent to which our findings can be extrapolated to situations involving non-binary states.

[5]It is equivalent to choosing an order but having the option to tell a receiver to wait.

1.1 Literature Review

Our paper contributes to the literature on information design and Bayesian persuasion. Aumann et al. (1995), Rayo and Segal (2010), Kamenica and Gentzkow (2011), and Bergemann and Morris (2016a,b) are some of the first models in this field. We apply this framework to settings where the sender persuades multiple receivers to take irreversible actions, allowing for information to be provided sequentially. In that way, we contribute to two areas: persuasion of multiple receivers and dynamic persuasion.

The literature on multiple receivers is mostly concerned with strategic interaction between those agents: they play a game in which payoffs depend on the underlying state of nature and the profile of actions of all agents. See the survey by Bergemann and Morris (2019) for a review of that literature. In contrast, in most of our paper, we shut down any strategic interactions between the receivers: each of them cares only about their "own" state and not at all about the actions of others. The only thing that links the receivers is that the states are correlated and hence a public persuasion of one receiver constrains the sender in their ability to persuade other receivers.

We define first-best payoffs based on what can be achieved with private communication. So, one way to interpret our results is that they show how the cost of publicity of communication can be mitigated or even completely removed with the appropriate dynamic persuasion. The restriction to public messages is also present in Alonso and Câmara (2016). That paper studies the simultaneous persuasion of multiple voters.

Our results are related to Bergemann and Morris (2016b) who show that if the receiver has access to additional information, the set of feasible outcomes is reduced. The connection to our paper is that since the sender provides information about multiple correlated states, information released about state one becomes a source of additional information about other states and vice versa. Our model shares the information leakage problem with this paper. Unlike in Bergemann and Morris (2016b), our additional information is chosen optimally by the sender, and the sender may decide not to follow optimal persuasion for one receiver to limit the indirect effect he creates for other receivers.

Our study of sequential persuasion procedures also contributes to the literature on dynamic persuasion, such as Hörner and Skrzypacz (2016) and Ely (2017). In that area, our work is particularly related to Ely and Szydlowski (2020) and Orlov et al. (2020).[6] These papers analyze the provision of information in a dynamic setting where the decision-maker's actions are irreversible. In those papers, the receiver decides in every period whether to stop or continue (stop putting effort or take an irreversible decision to remove a product from a

[6] See also Smolin (2021) who analyzes optimal evaluation policies for an agent who decides when to quit.

market). The decisions to stop or continue are not reversible because the receiver cannot go back in time. In our model, decisions taken by earlier receivers are also irreversible. Our discussion of non-additive preferences in the last section is also related to the bank run example in Ely (2017). That example illustrates the optimal private persuasion of multiple receivers and non-additive preferences of the sender. We discuss the benefits of sequential public communication.

There are three main differences between our model and those two models. Most importantly, in our model, the set of actions available to the receivers in later steps does not depend on the actions taken previously. In contrast, in those papers, if the receiver stops in time t, he has no other choices to make at later times. Second, we do not allow the sender to condition the past signals on the previous actions taken by the receivers. In our model, providing information is not used as an incentive. In Ely and Szydlowski (2020) and Orlov et al. (2020), a major part of the intuition of why gradual information release helps the sender is precisely the carrot aspect of future information. The sender can entice the receiver to delay stopping by promising additional information in the future. In contrast, since in our model, each receiver acts only once, they always make myopically optimal decisions. That allows a sharp characterization of how sequential information disclosure can help the sender reduce the cost of public information disclosure.[7]

There are also papers on sequential persuasion by multiple senders (see, for example, Board and Lu (2018) or Li and Norman (2021)). Those papers study strategic interactions among senders. Thus, in those papers, the sender's ability to provide information is constrained by the presence of other senders whose preferences are not necessarily aligned. In contrast, in our paper, the sender's ability to design information is constrained because the states of the world relevant for different receivers are correlated, and unlike in papers with multiple senders, our sender internalizes the effects of early disclosure on later persuasion.

Our paper is also related to the literature on multi-product firms (Gamp (2019)) and products with multiple characteristics (Turlo (2019)), where the sender determines the order for information acquisition. In our paper, the sender decides the order in which the receivers take action and what information they receive.

Finally, our paper is related to Arieli and Babichenko (2019), who study a sender with non-linear preferences. They examine a simultaneous procedure with multiple receivers but focus on private communication.

[7] Two other differences between our model and those two papers is that we have no discounting/delay costs and that we allow the sender to choose the order of actions, while in those papers, the order is determined by the passage of time.

2 The Model

There are $n+1$ agents. Agent $i = 0$ is a sender, and agents $i = 1...n$ are receivers. There are n binary states $\{\omega_i\}_{i=1}^n$, where $\omega_i \in \{L, H\}$. The common prior belief regarding these states does not need to be symmetric or independent. We denote by x_i the prior probability of $\omega_i = H$. When the receivers observe a public signal S (that can be correlated with all or some states), they update their beliefs about the n states. We let $x_{s,i}$ be the updated belief that $\omega_i = H$ conditional on observing $S = s$.

We assume that each receiver takes one action. In particular, receiver i takes action a_i that depends on his belief about ω_i. We denote this dependence by the function $a_i(x)$ for any posterior belief $x = x_{s,i}$.[8]

The sender's utility as a function of the action of receiver i is $f_i(a_i)$. The sender's utility as a function of the action profile, $\{a_i\}_{i=1}^n$, is additive in the different actions and is given by $\sum_{i=1}^n f_i(a_i)$. With some abuse of notation we also refer to f_i as a function of $x_{s,i}$; that is, $f_i(x_{s,i}) \equiv f_i(a_i(x_{s,i}))$. If the sender's payoff also depends on the state, then the payoff is $f_i(a_i, \omega_i)$. In that case, we define $f_i(x_{s,i}) \equiv E f_i(a_i(x_{s,i}), \omega_i)$, where the expectation is taken over ω_i with the belief $x_{s,i}$. Finally, we let $cav f$ denote the concavification of function f.

We compare a simultaneous (persuasion) procedure to sequential ones. An important assumption about the sequential procedures is that once receiver i takes his action, he cannot change it later even if new information comes to light and he regrets his earlier decision.

Definition 1. *A simultaneous procedure is described by a single public signal S that induces a vector of posterior beliefs $x_{s,i}$. For that procedure, the sender's expected payoff is:*

$$E_S[\sum_{i=1}^n f_i(x_{s,i})].$$

Definition 2. *A sequential procedure is described by a sequence of public signals $\{S_i\}_{i=1}^n$ and a permutation $\pi : \{1..n\} \to \{1..n\}$ specifying the order in which receivers take actions. In step i, the sender sends signal S_i, and receiver $\pi(i)$ takes action $a_{\pi(i)}$ (as a function of the posterior belief generated by all the signal realizations observed so far). The signal S_i may depend on the realizations of earlier signals, $\{s_j\}_{j=1}^{i-1}$.*

In the sequential procedure defined above, the permutation (the order in which the receivers make their decisions) is chosen ex-ante. In Section 4, we analyze a dynamic procedure

[8] Since we are focusing on the sender's payoffs, we do not specify the receiver's payoffs that determine which action is optimal for each belief. How the optimal actions change with beliefs depends on the receiver's joint preferences over actions and states. As usual in the models of Bayesian persuasion, when the agent has more than one optimal action we assume that ties are broken to the benefit of the sender.

that we refer to as a flexible sequential procedure (FSP). In this procedure, the sender can adjust the order of receivers based on the value of realized signals.

Definition 3. *A flexible sequential procedure (FSP) is a sequence of k signals* $\{S_i\}_{i=1}^k$ *and a conditional order of receivers* $\pi(S_1...S_i) \in \{0, 1, ...n\}$*, that specifies the sequence in which receivers take their actions conditional on the signals realized so far. In step i, conditional on the realizations of signals* $S_1, ...S_{i-1}$*, the sender sends signal* S_i *and the receiver* $\pi(S_1...S_i)$ *takes action* $a_{\pi(S_1...S_i)}$*. When* $\pi(S_1...S_i) = 0$*, no receiver takes an action in step i. The number of steps, k, can be random but is almost surely finite.*

Implicit in this definition is that each receiver must eventually be chosen for an FSP to be well-defined. That is, for every possible sequence of signal realizations, the conditional order $\pi(S_1...S_k)$ has to select each receiver once. Also, note that we allow for $k \geq n$. That is, we allow the sender to choose a procedure where no receiver takes action in some rounds.[9]

FSP offers greater flexibility, but this flexibility may not always be feasible. For instance, consider a company seeking regulatory approvals to market its product across various jurisdictions (e.g., states or countries). Before initiating the process of collecting and disclosing evidence to persuade regulators, the company needs to determine the sequence in which it will apply for permissions. Nevertheless, the FSP not only provides a theoretical benchmark, it also holds real-world relevance. For example, as mentioned in the Introduction, it is a model of a situation where information continuously arrives through some external process. Beliefs adjust continuously, and eventually, the states become revealed. The sender primarily determines the timing of the receivers' actions. At each point in time, the sender decides whether to prompt a receiver to take action and, if so, whom to prompt.

Finally, in some applications, the order of receivers may be beyond the sender's control. It may be fixed exogenously. In that case, we should interpret our result for the sequential procedure as showing that at least one order of receivers exists for which the sequential procedure strictly improves the sender's payoff.

Remark 1. *We interpret the game as having multiple receivers taking a single action each. Equivalently, we could allow some of the actions to be taken by the same receiver or even to have just one receiver making all of the decisions. In the latter case, we assume that after action* a_i *the receiver does not learn additional information about the other states beyond what is revealed by the sender. For example, the receiver commits to an action, but the*

[9] Another type of procedure with an intermediate degree of flexibility would be to allow the sender to choose the first receiver, disclose the first signal, make that receiver act, and then, depending on the realized signal, pick the next receiver. The substantial difference is that FSP either allows recommending the receiver to wait or allows choosing each receiver after the signal in a given round is realized. The payoffs of that alternative persuasion procedure would be in between those of the sequential procedure and the FSP.

payoff consequences are delayed until all actions are chosen. This interpretation captures situations where the receivers represent a voter (or a set of voters as in a budget or hiring committee), and the sender chooses a sequence of proposals to present to them. The time between votes on the different issues can be much smaller than the time needed to learn the consequences of the already made decisions.

2.1 Preliminary Result: First Best

We start by establishing an upper bound on the sender's payoffs:

Lemma 1. *For any procedure (simultaneous or sequential) the sender's payoff is at most*

$$FB(\{x_i\}) \equiv \sum_{i=1}^{n} cav f_i(x_i).$$

As we know from Kamenica and Gentzkow (2011), $cav f_i(x_i)$ equals the optimal payoff that the sender could obtain from action i if it was the only action. Hence, FB, which stands for first best, represents the sum of payoffs if the sender could optimize each dimension separately. It is an upper bound for any procedure, sequential or simultaneous, since no procedure can achieve for dimension i a payoff greater than $cav f_i(x_i)$.

Note that this payoff would be obtained by the optimal simultaneous procedure if the sender could communicate privately with every receiver. With private communication, there is no information leakage: when the sender persuades receiver i about state ω_i, he can do it without changing the beliefs about the other states. In contrast, providing public information about state ω_i reveals information about other states as well. This information leakage constrains the sender's ability to persuade optimally in every dimension at once. Therefore, when the correlation of states is strong enough, the sender may not be able to achieve the first-best payoff.

3 Simultaneous versus Sequential Persuasion

3.1 Strict Improvement from Sequential Procedure

Our goal in this section is to compare simultaneous to sequential procedures. Note that the outcome of any simultaneous procedure can be replicated by a sequential procedure with all signals except for the first one being uninformative. So, trivially the optimal sequential procedure weakly dominates the simultaneous one. We characterize when the improvement is strict.

Our main result is that if there exists a scope for improvement (i.e., the simultaneous procedure cannot achieve the first-best bound), then the optimal sequential procedure strictly improves upon the simultaneous one. An interesting fact is that the sequential procedure may also fail to achieve the first-best bound. We show this in the following subsection, where we present a few examples. We argue that:

Theorem 1. *The optimal sequential procedure strictly dominates (in terms of the sender's payoffs) the optimal simultaneous procedure if and only if the latter fails to achieve the first-best payoff, $FB(\{x_i\})$.*

The proof follows a simple intuition. Consider the optimal simultaneous procedure and assume it fails to achieve the first-best payoff. If this is due to a receiver being provided "too much" information one can benefit from making him make the decision first, and providing him with less information. If this is due to a receiver being provided "too little" information, one can benefit from making him make the decision last, and providing him with additional information.

Proof. The 'only if' direction follows from the fact that $FB(\{x_i\})$ is an upper bound on both procedures. So we only need to prove the 'if' part. Consider a simultaneous procedure with signal S that fails to achieve the first-best outcome, $FB(\{x_i\})$. This implies that for at least one receiver, i:

$$E_S[f_i(x_{s,i})] < cav f_i(x_i).$$

We shall argue that a sequential procedure exists that improves the payoff from action i while keeping the payoffs from all other actions unchanged. Our proof is based on considering two cases that are illustrated in Figure 1.

We begin by assuming that $cav f_i(x_i) > f_i(x_i)$. Based on this, there exist $x_i^L, x_i^R \in [0,1]$ such that $x_i^L < x_i < x_i^R$ and payoffs satisfy $cav f_i(x_i^L) = f_i(x_i^L)$, $cav f_i(x_i^R) = f_i(x_i^R)$, $x_i = \lambda \cdot x_i^L + (1-\lambda) \cdot x_i^R$, and $cav f_i(x_i) = \lambda \cdot f_i(x_i^L) + (1-\lambda) \cdot f_i(x_i^R)$. Moreover, for any $x_i \in (x_i^L, x_i^R)$, $f_i(x_i) < cav f_i(x_i)$.[10]

- Case 1: For some realizations $S = s$, we have $x_{s,i} \in (x_i^L, x_i^R)$.
  Since $f_i(x_{s,i}) < cav f_i(x_{s,i})$ we construct a sequential procedure where receiver i takes his action last. All other receivers take actions based on the original signal S (in an arbitrary order). In the last step, conditional on observing $S = s$ the procedure provides more information regarding state i, which induces posteriors of x_i^L and x_i^R for state ω_i. This is a strict improvement.

[10] Generically, x_i^L, x_i^R are unique. In case they are not, we define them to be the closest points to x_i.

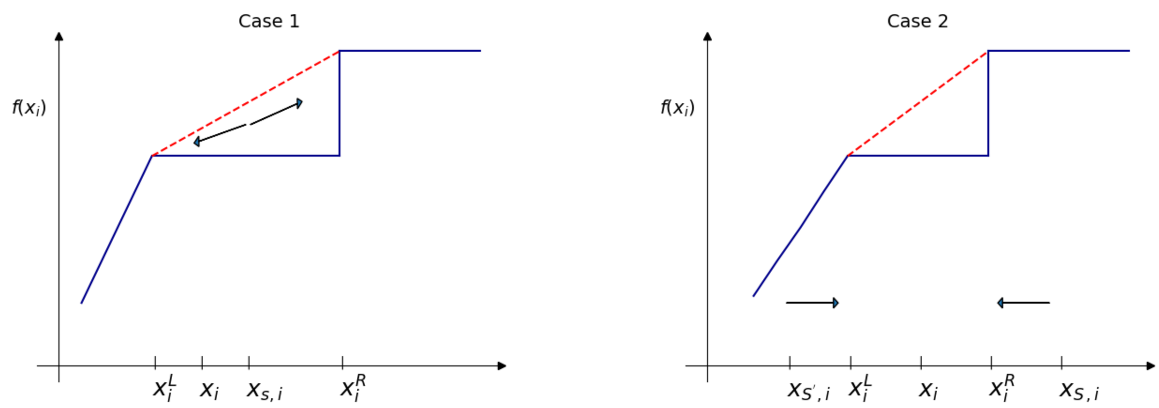

Figure 1: Proof of Theorem 1

- Case 2: For all realizations $S = s$ we have $x_{s,i} \notin (x_i^L, x_i^R)$.
  In the sequential procedure, receiver i takes his action first. In the first stage, the sender reveals signal S', which is a garbling of S. By Lemma 3 in the Appendix, there exists a garbling S' of S which has two realizations: s_1', s_2' so that $x_{s_1{}',i} = x_i^L$ and $x_{s_2{}',i} = x_i^R$. After S' is realized and the receiver i takes his action, the original S is revealed in the second step, and the remaining receivers take their actions (in an arbitrary order). While the original signal does not achieve $cav f_i(x_i)$, this new sequential procedure does and leaves the payoffs for all other dimensions unchanged, so it is also a strict improvement.

Finally consider the case when $f_i(x_i) = cav f_i(x_i)$. In this case, the expected payoff from action i is strictly lower when the signal S is revealed. Hence, one can improve if, in the first step, receiver i takes action a_i based solely on the prior. In the second step, signal S (from the simultaneous procedure) is revealed (and no more information is revealed later). Then the remaining receivers take their actions in an arbitrary order.[11]

□

[11] One can view this as a special case of Case 2 when we have $x_i^L = x_i^R$.

3.2 Examples

To illustrate Theorem 1, we consider three simple examples. In these examples, a sender wishes to affect the binary actions of two receivers: $i = 1, 2$. Receiver i selects $a_i = 1$ provided that his posterior belief $x_{s,i}$ exceeds a certain threshold and chooses $a_i = 0$ otherwise. We assume that $f(a) = a$ so the sender maximizes $E(\sum_i a_i)$.

Example 1. *The thresholds for both receivers are* 0.8. *The two states are perfectly negatively correlated. That is, the prior is that exactly one of the two states is* H. *The prior is that each state is* H *with probability* 0.5.

Consider first a simultaneous procedure. If the sender reveals no information, both receivers pick actions based on the prior and choose $a_i = 0$. *This implies that the sender's payoff is zero. If, instead, the sender provides full information, exactly one receiver picks* $a_i = 1$, *and the sender's payoff is one. This is an optimal simultaneous procedure as at most one receiver would pick* $a_i = 1$.[12] *For example, if receiver* 1 *selects* $a_1 = 1$, *then* $x_{s,1} = Pr(\omega_1 = H) \geq 0.8$. *This implies that* $x_{s,2} = Pr(\omega_2 = H) \leq 0.2$ *and receiver* 2 *picks* $a_2 = 0$.

Consider now a sequential procedure. The sender first provides a signal about ω_1 *that induces a posterior belief of either* 0 *or* 0.8 *with probabilities* $3/8$ *and* $5/8$, *respectively. Conditional on the posterior of* 0, *receiver* 1 *selects* $a_1 = 0$ *and receiver* 2 *selects* $a_2 = 1$ *in the second step. Conditional on the posterior of* 0.8, *receiver* 1 *selects* $a_1 = 1$ *in the first step. At the end of step one, receiver* 2 *assigns a probability of* 0.2 *that* $\omega_2 = H$. *In the second step, the sender can provide more information. With a positive probability, the sender can convince receiver* 2 *to pick* $a_2 = 1$. *In particular, given a prior of* 0.2, *the sender can induce a posterior of* 0.8 *with probability* $1/4$. *Overall, this sequential procedure induces at least one receiver to pick* $a_i = 1$, *and sometimes both receivers take the action* $a_1 = a_2 = 1$. *The sender's expected payoff from this sequential procedure is* $3/8 + 5/8 * (1 + 1/4) = 1.156$.[13]

Example 2. *Similar to example 1, but the thresholds for receivers* $i = 1, 2$ *are* 0.6 *and* 0.8, *respectively. In addition, the two states are perfectly positively correlated. The prior is given by* $Pr(\omega_1 = \omega_2 = H) = Pr(\omega_1 = \omega_2 = L) = 0.5$.

Consider first the simultaneous procedure. Given the perfect correlation, the sender's payoff is a function of the belief regarding the common state. If the probability that it is H *is lower than* 0.6, *then the sender's payoff is zero. If it is between* 0.6 *and* 0.8, *it is one. If it exceeds* 0.8, *then it equals* 2. *Given the prior of* 0.5, *using a standard concavification*

[12] This optimal procedure is not unique. For example, information that induces posteriors of 0.9 and 0 yields the same payoff.

[13] We prove in the Appendix that this is the best sequential procedure in this example.

argument we conclude that the optimal procedure is to generate a posterior of 0.8 *with a probability* 5/8 *and zero with a probability of* 3/8*. The sender's payoff is then* $2 * 5/8 = 1.25$.

Consider now a sequential procedure in which the sender first wishes to convince the first receiver to take action $a_1 = 1$*. In the first step, the sender induces a posterior of* 0.6 *with a probability* 5/6 *and a posterior of* 0 *with a probability* 1/6*. His expected payoff in the first step is* 5/6*. In the second step, if he induced a zero posterior, his payoff is zero. If he has induced an interim posterior of* 0.6 *then he can induce a final posterior of* 0.8 *with a probability of* $6/8 = 3/4$ *(and a posterior of* 0 *otherwise). Hence, his expected payoff from the second receiver is* $(3/4) * (5/6) = 5/8$*. The overall sender's expected payoff is* $5/6 + 5/8 = 1.458$.

Example 3. *Same as in example 2, but the two states are independent. That is, each state is H with a probability of* 0.5*, and these events are independent. Consider a simultaneous procedure in which the sender induces a posterior for* ω_1 *of* 0.6 *with probability* 5/6 *and zero with probability* 1/6*. Given the independence, the sender can simultaneously induce a posterior for* ω_2 *of* 0.8 *with probability* 5/8 *and zero with probability* 3/8*. The first receiver selects* $a_1 = 1$ *with a probability* 5/6*, and the second receiver selects* $a_2 = 1$ *with a probability* 5/8*. This leads to a total payoff of* 1.458*. In this case, no sequential procedure can improve the sender's payoff.*

Examples 1 and 2 highlight a key reason why a sequential procedure generates a higher payoff for the sender: the first receiver sometimes regrets his decision once he sees the public signal in the second step. A key difference between the two examples is that in the second example, the sender achieves the first-best outcome in the sequential procedure. In the first example, he does not. The sender in example 3 achieves the first best with both procedures.

The three examples combined illustrate the result in Theorem 1. The optimal sequential procedure strictly dominates the optimal simultaneous one if and only if the simultaneous procedure fails to achieve the first-best payoff (the upper bound). Table 1 summarizes that point:

	Simultaneous	Sequential	First-Best
Example 1	1	1.156	1.25
Example 2	1.25	1.458	1.458
Example 3	1.458	1.458	1.458

Table 1: Payoffs in the Three Examples

It is important to note that: (i) this result is not based on the optimal sequential procedure always achieving the first-best payoffs, as can be seen in Example 1, and (ii) this result applies to a general action space and general payoffs.

For two receivers with binary decisions, we provide additional analysis in the Appendix. One may ask, under what conditions does the simultaneous procedure fail to achieve the first-best outcome? This is when the sequential procedure achieves a higher payoff. In the Appendix, we examine this question in a setup based on the above examples. We show that this would occur if and only if there is sufficient negative correlation between the two states.

4 Flexible Sequential Persuasion

In the previous section, we assumed that the sender predetermines the order of receivers in a sequential procedure. This section considers "flexible sequential" procedures. As we discussed above, this would fit a different scenario in which it is not the sender who generates new information. The sender's role primarily involves timing the actions of the receivers. At each point in time, the sender determines whether to prompt a receiver to take action and, if so, whom to prompt.

To gain some intuition, consider Example 1 from the previous section. The sender can achieve the first-best payoff by employing the following strategy. In the first step, the sender induces a posterior belief $Pr(\omega_1 = H)$ of 0.8 or 0.2. Given the perfect negative correlation, the signal induces a posterior belief $Pr(\omega_2 = H)$ of 0.2 and 0.8, respectively. The sender asks receiver i to take action a_i provided that $Pr(\omega_i = H) = 0.8$. Without loss of generality, suppose that this is $i = 1$. Given that $Pr(\omega_1 = H) = 0.8$ the receiver selects $a_1 = 1$. Note that the belief is now that $Pr(\omega_2 = H) = 0.2$. In the second step, the sender follows standard concavification. He induces a posterior of 0.8 with a probability of 0.25 and a posterior of 0 with a probability 0.75. This implies that receiver $i = 2$ selects $a_2 = 1$ with a probability of 0.25. The sender's expected payoff equals $E(a_1 + a_2) = 1.25$, which is the first-best payoff.

As we shall show next, this illustrates a more general result. An optimal FSP, which we name the 'Pacman procedure', always achieves the first-best payoff.

4.1 The Pacman procedure

We now characterize an optimal FSP as follows: We shall assume that for all i we have that $cav f_i(x_i) > f_i(x_i)$. If this is not the case, and $cav f_i(x_i) = f_i(x_i)$, then the optimal FSP starts with receiver i, and this receiver is asked to take action i without receiving any information.

Recall from the proof of Theorem 1 that x_i^L, x_i^R denote the optimal concavification beliefs for action a_i (they depend on the prior, but we suppress that notation). Let $x_i^L < x_i < x_i^R$. The Pacman procedure that achieves the first-best payoff is based on a sequence of signals and a contingent sequence of actions with the following property. The receiver takes action

i only when $x_{s,i} \in \{x_i^L, x_i^R\}$ (we abuse notation by writing $x_{s,i}$ as the posterior belief given the realized sequence of signals). The idea is to construct a dynamic Bayesian persuasion procedure where, as we disclose information about some states, the beliefs regarding all states ω_i remain in every step in the intervals $[x_i^L, x_i^R]$. When the belief hits one of the boundaries of these intervals, receiver i is asked to take action and we remove that receiver from the continuation problem. This ensures that the procedure achieves the first-best payoff, provided that it ends in a finite time because from the ex-ante perspective, beliefs are a martingale, and all receivers take actions exactly at their concavification points.

For any i, consider some signal S that is correlated with ω_i, and conditional on ω_i is independent of any other ω_j. Suppose that the signal S is rich in the sense that for every $p \in [0,1]$ there exists s such that $x_{s,i} = p$. Finally for that signal define $h_{i,j}(p) \equiv Pr(\omega_j = H | x_{s,i} = p)$. This function describes how sensitive the beliefs about state ω_j are in response to signals about state ω_i. The following technical lemma shows that this sensitivity is at most one:

Lemma 2. $|h'_{i,j}(p)| \leq 1$.

To construct an optimal FSP (that achieves the first-best payoff FB), define:

$$\Delta_i = \min\{x_i^R - x_i, x_i - x_i^L\}$$
$$\Delta = \min_i \Delta_i$$
$$i^* = \arg\min_i \Delta_i.$$

Without loss of generality, assume that the minimization in Δ_{i^*} is achieved by $x_{i^*}^R$ (that is, $x_{i^*}^R - x_{i^*} \leq x_{i^*} - x_{i^*}^L$).

We now construct the Pacman procedure. In step 1, we choose a signal conditional on ω_{i^*} (and conditionally independent of the other states) that generates posterior beliefs about ω_{i^*} of $x_{i^*} \pm \Delta$, each with probability 0.5. Consider the posteriors for all other states. Lemma 2 implies that for both realizations of the signal in step 1:

$$\forall i, s : x_{s,i} \in [x_i^L, x_i^R].$$

Conditional on the posterior for ω_{i^*} being $x_{i^*} + \Delta = x_{i^*}^R$, we ask the receiver i^* to act; otherwise, we tell all receivers to wait. We then iterate on this procedure. If receiver i^* acts, we remove it from the consideration set. Then, for the posterior beliefs from the previous step, we redefine Δ and i^* and repeat.

To see that this procedure achieves the first-best payoff, note that in each step, one of the receivers acts with probability 0.5, and actions are taken only at the optimal concavification thresholds. This implies that the procedure ends in finite time almost surely. This proves the following theorem: [14]

Theorem 2. *The Pacman procedure achieves the first-best payoff.*

Finally, the Pacman procedure helps us understand an optimal sequence of persuasion. With this procedure, it is optimal to rank decisions based on the distance between the prior and the nearest concavification threshold (the Δ_i's) and then persuade based on this order. The only caveat is that the ranking of Δ_i's can change due to the disclosed information, so the order of persuasion depends not only on the ex-ante parameters but also on the information learned in the process of persuasion. The other takeaway from this construction is that with a flexible procedure, the order of information matters less than disclosing information gradually and then choosing the proper order of actions. The most important thing about this procedure is that information is disclosed gradually to ensure that a concavification threshold is not jumped over in any step.

5 Simple versus Flexible Sequential Persuasion and Optimal Order of Persuasion.

We now turn to when the optimal FSP can strictly improve upon the simple sequential procedure we examined before. Since the FSP always achieves the first-best payoff, the question is under what conditions the simple sequential procedure can achieve it too. In order to provide a partial answer to this question, we start with the special case of perfectly correlated states. That was the case in Examples 1 and 2. In general, perfect correlation makes it the hardest to achieve the first-best payoff.

5.1 Perfectly Correlated States

Let x denote the (common) prior belief that all states are $\omega_i = H$.[15] The payoff of the optimal simultaneous procedure is:

[14] While we construct the Pacman procedure to take more rounds than the number of receivers, it is possible to reduce the number of steps. Namely, we can create a more complex sequence of signals to take n steps and achieve the first-best payoff exactly. The intuition is that when the first signal realization moves the beliefs to x_{i*}^L, the sender does not need to reveal the signal. Instead, he can automatically compute the next optimal signal and reveal it jointly with the first signal realization. If the realization of the second signal also does not move beliefs to any of the concavification thresholds, we iterate. In the limit, we obtain a random variable such that every realization moves the beliefs to at least one of the optimal concavification thresholds without moving any beliefs outside the $[x_i^L, x_i^R]$ region. So in every step, one of the actions is taken, and information leakage never moves beliefs 'too far.'

[15] Without loss of generality, we assume perfect *positive* correlation. If some states are perfectly negatively correlated with ω_1, then we switch the definition of those states, inducing a positive correlation.

$$U^*_{Sim}(x) = cav(\sum_{i=1}^{n} f_i)(x). \tag{1}$$

The payoff from a sequential procedure with two receivers acting in the order $(1, 2)$ is $cav(f_1 + cav f_2)(x)$. The intuition is that if the sender induces in the first stage posterior $\hat{x}$, in the second stage, the persuasion payoff is $cav f_2(\hat{x})$. Taking this into account, the sender realizes that the total payoff as a function of the first-period posterior is $(f_1 + cav f_2)(\hat{x})$. This, in turn, is maximized in stage 1 by the standard concavification at the prior, yielding $cav(f_1 + cav f_2)(x)$.

For more receivers and some permutation $\pi : \{1..n\} \rightarrow \{1..n\}$, the optimal sequential procedure that follows this order yields a payoff that can be defined recursively:

$$H_n(\pi) \equiv cav f_{\pi(n)}$$
$$H_i(\pi) \equiv cav(f_{\pi(i)} + H_{i+1}(\pi)).$$

The payoff of the optimal sequential procedure given prior x can therefore be expressed as:

$$U^*_{Seq}(x) = max_\pi H_1(\pi)(x). \tag{2}$$

The result in Theorem 1 in the case of perfect correlation of states can be viewed as the comparison of (1) to (2). In particular, whenever $U^*_{Sim}(x) < FB(x)$ then $U^*_{Sim}(x) < U^*_{Seq}(x)$.

This characterization can be used to show that the sequential procedure in Example 1 is optimal despite not achieving the first-best payoff (see the Appendix).

Our goal in this section is to examine when the simple sequential procedure achieves the first-best payoff. Define

$$\Delta_i^L \equiv x - x_i^L$$
$$\Delta_i^R \equiv x_i^R - x,$$

where recall that x_i^L and x_i^R are the optimal concavification thresholds to the left and to the right of the prior for receiver i (given the prior). Let π^L be the list of $i's$ in the order of Δ_i^L and analogously define π^R. For example if $n = 3$ and the vectors are $\Delta^L = (0.1, 0.3, 0.2)$ and $\Delta^R = (0.4, 0.1, 0.2)$, then $\pi^L = (1, 3, 2)$ and $\pi^R = (2, 3, 1)$. In words, π^L and π^R are the orderings of the receivers in terms of the ranking of the optimal concavification thresholds to the left and to the right of the prior, respectively. Note that if there are ties in some Δ_i's (see Example 2), these orderings are not unique (all ways of resolving these ties are allowed).

Proposition 1. *If states are perfectly (positively) correlated, the optimal sequential procedure achieves the first-best payoff if and only if there exist orderings such that $\pi^L = \pi^R$ (i.e., for at least one resolution of ties).*

When this condition is satisfied, the optimal sequence of persuasion follows these orderings: the sender persuades first the receiver with the smallest Δ_i^R, then the second smallest, and so on.

Proof. Suppose this condition is satisfied. That means that the concavification thresholds are nested around the prior. Moving in the proposed sequence, in every step, we reach one of the two concavification thresholds for one of the receivers, eliminating them deterministically one by one. Specifically, let $i = \pi^L(1)$. In step 1 choose a binary signal S_1 inducing posteriors $x_s \in \{x_i^L, x_i^R\}$. Let receiver i act and remove him from the consideration set. Next, redefine $i = \pi^L(2)$ and again choose a binary signal that induces posteriors $x_s \in \{x_i^L, x_i^R\}$. Since the thresholds are nested, in step $k+1$ the posteriors remain interior to $[x_j^L, x_j^R]$ for every receiver j later in the sequence than receiver $\pi^L(k)$. That allows the sender to induce posterior beliefs equal to the corresponding optimal thresholds in every step, proving that this order achieves the first-best payoff.

If the two orderings are not the same, then this implies that there are two receivers i, j such that $\Delta_i^L < \Delta_j^L$ but $\Delta_i^R > \Delta_j^R$ (as in Example 1). In that case, it is not possible to achieve the first-best payoff with the simple sequential procedure, since if the sender chooses an order of receivers such that the receiver i acts before receiver j (without loss of generality), then to achieve the first-best payoff for receiver i we need to induce with positive probability the belief $x_i^R > x_j^R$. That, in turn, means that we do not obtain $cav f_j(x)$ on receiver j. □

5.2 Imperfectly Correlated States

Proposition 1 allows us to also provide intuition about the imperfectly correlated case. Given priors x_i generalize our previous definition:

$$\Delta_i^L \equiv x_i - x_i^L$$
$$\Delta_i^R \equiv x_i^R - x_i,$$

and define the orderings π^L, π^R as before.

First, suppose that all the states are positively correlated (in the sense that all $h'_{i,j}(p)$ described in Lemma 2 are non-negative). If the orderings of dimensions in terms of ranking of Δ_i^L and Δ_i^R are the same ($\pi^L = \pi^R$) defined in the proof of Lemma 2, then the sequential procedure achieves the first-best payoff for all priors and (positive) correlations. Furthermore, if the orderings are not the same, no sequential procedure can reach the first-best payoff for a sufficiently high correlation. The intuition is the same as in the proof of Proposition 1. We can start with the receiver with the smallest Δ_i and induce the concavification thresholds. Since the $\Delta_i's$ are nested, we do not "jump over" the concavification thresholds for other receivers even with perfect correlation. When correlation is imperfect, the posteriors on the other dimensions are even closer to the starting priors.

Second, when the correlations are negative, having the receivers ranked in the same order in both directions is not enough. Note that if there are more than two receivers and a general imperfect negative correlation, it may be impossible to change the definition of states to ensure that all pairwise correlations are positive. The reason is that when we move x_i to x_i^R, even though we are guaranteed not to overshoot any x_j^R, we may overshoot some x_j^L. In that case, a sufficient condition for the sequential procedure to achieve the first-best payoff for all correlations is that for all i, j, if $\Delta_i^R < \Delta_j^R$, then both $\Delta_i^L < \Delta_j^L$ and $\Delta_i^R < \Delta_j^L$. That condition is also tight in the following sense. If one of the inequalities is reversed, for some correlation structure (positive or negative), the optimal sequential procedure does not achieve the first-best payoff.

6 Persuading Acceptance

In this section, we specialize the model so that receivers make binary decisions - accept or reject. This allows us to obtain additional results and relax some of the assumptions we have made before. We assume that the receivers' payoffs are such that there is a closed convex set of posterior beliefs under which that receiver accepts (and rejects for the complement of those beliefs). Further, we assume that for every receiver there is a state that, if revealed, he accepts. Note that while it is more restrictive than what we have considered before, all the examples we have analyzed satisfy these more restrictive assumptions.

We begin by characterizing the optimal sequence of persuasion. We then consider non-additive payoffs induced by voting on a common decision by a committee. Finally, we partially extend Theorem 1 beyond binary states and explain why Theorem 2 does not generically extend.

6.1 Optimal Sequence of Persuasion

Recall from the previous section the definition of optimal concavification thresholds x_i^L and x_i^R. In the case of persuasion to accept, it must be that either $x_i^L = 0$ or $x_i^R = 1$ (or both are equal to the prior if it is already in the acceptance set).

Consider the following notion of alignment between the receivers:

Definition 4. *The receivers are* aligned *if*

1. *States are positively correlated: For any i, if the information is revealed about state i that conditional on ω_i is independent of the other states, and the realized signal increases the posterior belief $x_{s,i}$, then all the other beliefs $x_{s,-i}$ weakly increase.*

2. *For each receiver the acceptance set includes state $\omega_i = H$.*

A special case of alignment is when the states are perfectly correlated. Naturally, the receivers then have a common prior and vary only in the smallest belief, x_i^R, necessary for receiver i to accept.

In the case of accept/reject decisions of aligned receivers, we get the following full characterization of the optimal sequential procedure, which is a corollary of Proposition 1 and the result in Section 5.2:

Corollary 1. *If receivers make accept/reject decisions and are aligned, then the optimal sequential procedure ranks the receivers based on $\Delta_i^R = x_i^R - x_i$ and achieves the first-best payoff.*

6.2 Non-additive payoffs - Persuading Voters

A natural application for persuasion of multiple receivers making accept/reject decisions is the persuasion of voters or committee members. However, since committee decisions are based on a certain majority, this requires that we consider non-additive payoffs.

Consider first the case of a unanimous vote. The sender's payoff is one if all receivers approve and zero otherwise. The receivers' payoff is zero if the vote fails and the payoff is state-dependent if all receivers approve. Would the sender benefit from a sequential procedure in that case?

Recall Example 1, in which there is a perfect negative correlation. With a simultaneous procedure, the sender will never be able to convince both receivers to approve. But as we saw above, in the additive case, the sequential procedure results in a positive probability that both receivers approve. However, this result depended on the preferences of the receivers being independent of the other receivers' actions. Consider a sequential vote requiring unanimous approval. Let Agent 1 vote first. If Agent 1 would approve when his beliefs were above 0.8 he would not be forward-looking because this decision would ignore that later Agent 2 approves only if the beliefs drop below 0.2. If Agent 1 were forward-looking, he would condition his decision on the fact that his vote matters only in the event that later Agent 2 approves. If so, he should not approve even when his belief is above 0.8.

We argue that this observation generalizes:

Proposition 2. *Assuming forward-looking agents and a unanimous vote, the optimal sequential procedure leads to the same payoff as the optimal simultaneous one.*

Proof. Consider the optimal sequential mechanism. Without loss of generality, we can restrict attention to binary signals. Each agent accepts if and only if he observes a signal realization of one. We prove the claim by induction. Consider first the case of $n = 2$ agents. Agent 2's decision is relevant only when Agent 1 approves. Assuming forward-looking behavior, Agent 1's decision is based on the pivotal event in which Agent 2 approves; thus, Agent 1 never regrets his decision. This implies that if we reveal to Agent 1 the signal that Agent 2 obtains, he would behave the same. Hence, this simultaneous mechanism would lead to the same outcome. Assume it holds for n agents and consider $n + 1$. Given the induction assumption, we can restrict attention to a mechanism in which in the first step Agent 1 receives a signal and decides. In the second step, the remaining n agents receive a joint simultaneous signal $\in \{0, 1\}^n$, where $s_i = 1$ is a recommendation to approve and $s_i = 0$ to reject. Being forward-looking, Agent 1 conditions on the event that all future agents observe the signal vector 1^n. Hence, even if he observed the joint signal, he would not change his decision, and the mechanism could be made simultaneous. □

We next consider the case of a majority vote and argue that the sender may strictly prefer a sequential procedure.

Example 4. *Suppose that there are* 3 *agents, and the sender seeks approval by a majority. The states of agents* 1 *and* 2 *are perfectly correlated, and they are perfectly negatively correlated with the third state. The thresholds for approval for agents* $1, 2$, *and* 3 *are* $0.7, 0.8$, *and* 0.8, *respectively. The prior for all three states is* 0.5.

Claim 1. *Assuming forward-looking agents, in Example 4 a sequential procedure generates a higher sender's payoff than a simultaneous one.*

Proof. In the case of the optimal simultaneous procedure, the only case in which the sender is able to secure a majority vote is by convincing agents 1 and 2. Optimally, the sender must induce a posterior of 0.8 or zero for states 1 and 2. Both agents will approve when their posterior is 0.8 which occurs with a probability of 5/8.

In a sequential procedure, the sender will induce a posterior of 0 and 0.7 for states 1 and 2 and ask agent 1 to approve which will occur with a probability of 5/7. After convincing receiver 1, the sender induces a posterior of 0.8 and 0.2 for states 1 and 2, which implies a posterior of 0.2 and 0.8 for state 3 because of the perfect negative correlation. Thus, the sender can obtain approval either from Agent 2 or Agent 3 with certainty. Thus the sequential procedure does strictly better. □

6.3 Beyond Binary Case

In all our analysis so far we have assumed that each receiver cares only about a binary state. While that is a good assumption for some applications, it may be a restrictive assumption for others.

We now show that after (suitably generalized) alignment of the receivers, Theorem 1 can be extended to the multidimensional case, but Theorem 2 generically does not.

To study the case of the multi-dimensional state, we assume that, without loss of generality, there is one state ω that has K potential realizations ω_k with $k \in \{1, ..., K\}$. The receivers have a common prior x that is a vector with elements $x_k = \Pr\{\omega = \omega_k\}$ and without loss of generality we assume that the prior is fully mixed (assigns positive probability to all ω_k).

In the space of beliefs, each receiver has an acceptance region that contains at least one of the extreme beliefs (i.e., $x_k = 1$ for some k). We assume that the acceptance sets are convex and closed.

We extend the definition of alignment as follows:

Definition 5. *In the case of multidimensional states, the receivers are aligned and nested if*

- *Their acceptance sets contain the same set of extreme beliefs.*
- *The receivers can be ordered in a way that their acceptance sets are nested, with receiver 1 acceptance set containing the acceptance set of receiver 2, and so on.*

Moreover, the receivers are strictly nested if there exists $\epsilon > 0$ such that for any fully mixed beliefs on the boundary of any two receivers' acceptance sets, there is at least an ϵ distance between those beliefs.

An example with $k = 3$ and 2 receivers is shown in Figure 2. The vertices of the triangle represent the extreme beliefs. The orange triangle is the acceptance set of receiver 1, and the blue triangle is the acceptance region of receiver 2. The prior (without loss of generality) is outside both acceptance sets. Note that the sets are strictly nested - while they share parts of the boundaries with beliefs that rule out one of the states, for all fully mixed beliefs on the boundary of the orange set, there is a strictly positive distance to any belief in the blue set.

Proposition 3. *In the case of aligned and strictly nested receivers, if the prior is outside any of the acceptance sets:*
1) The optimal sequential procedure strictly improves upon the simultaneous procedure.
2) The optimal order of receivers in a sequential procedure coincides with the nesting order.
3) Generically, the sequential procedure does not achieve the first-best payoff.

We present the full proof in the appendix, but the intuition is as follows. For parts 1 and 2 the intuition is analogous to the intuitions we presented above for the binary states. The intuition for part 3 depends on the multidimensionality of the state.

First, the first-best payoff (with private messages) is achieved by finding for each receiver two posteriors, one on the (closure of) the fully mixed boundary of the acceptance set of that receiver and another one on the boundary of the rejection set. These two beliefs lie on the same line as the prior belief. Generically (part 3), these lines are unique and different for every receiver. Hence, it is impossible to replicate that outcome via any sequential procedure (or FSP) since achieving the first-best payoff for one of the receivers implies getting away from the optimal lines for other receivers. In the two-dimensional state case the lines are constrained to be in just one dimension, which explains the difference from Theorem 2.

Second (part 2), as before, in the optimal procedure we want receivers to accept with the least possible information because it increases the probability they would accept and reduces the amount of information leakage that constraints the sender's ability to persuade future receivers. This implies that in the optimal sequential procedure if receiver i were called to accept before receiver j with a larger acceptance set, then we are providing j with too much information. We could get the same payoff by changing the order of the receivers. Moreover, we could get a strict payoff improvement by garbling some information (combining the signal that recommended i to accept with some signals that recommended j to reject) to bring the posterior belief at the time j accepts closer to the boundary of his set, but still inside the set. That would increase the probability that j accepts without changing the behavior of the other receivers.

Third (part 1), the optimal simultaneous procedure either provides too little information by making only a subset of the receivers accept or too much information by persuading all of them to accept. In the first case, we can improve by a sequential procedure by fully revealing the state after the first set of receivers accepts (and thus getting additional receivers to accept with a positive probability). In the latter case (all accepting), we can improve by having the receivers that are easiest to persuade (i.e., those with the largest acceptance sets) to move first with less information.[16]

[16] In the statement of the Proposition we assumed that the prior is in the rejection set of every receiver. This can be generalized. First, if the prior is inside the smallest acceptance set, then no persuasion is necessary, and hence the simultaneous procedure achieves the first-best payoff. Second, if it is in the second-smallest acceptance set, then the sequential procedure achieves the first-best payoff by telling $n-1$ receivers to act first with no information and then doing the optimal persuasion of just one receiver. However, the simultaneous procedure does not achieve the first-best payoff. Finally, if the prior is outside the second-smallest acceptance set, all the statements in the proposition hold, just that the order in part 2 may not be uniquely optimal.

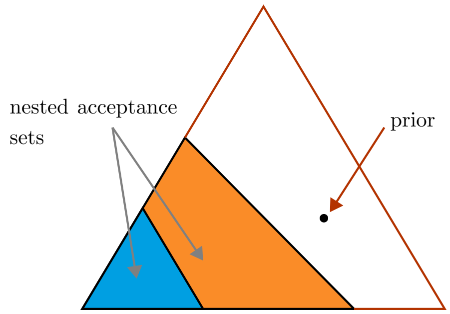

Figure 2: Strictly Nested Aligned Receivers.

7 Conclusion

In this paper, we have analyzed a sender who provides information to influence multiple receivers. We have shown that if states are correlated, the receiver may suffer from revealing information publicly, instead of sending receivers private messages. Our main result is that in the case of binary states and additive preferences, whenever public communication hurts the sender, he strictly benefits from sequential communication and decision-making by the receivers. We show how to construct such improving communication strategies in two cases - when the sender has to pre-commit to the sequence of actions and when he has even more flexibility and can adjust the order of receivers based on the realized information. In the latter case, the sender can recover full first-best payoffs. We finish the paper discussing a subclass of problems where receivers take binary actions: accept or reject. We show that with unanimous voting and forward-looking receivers, the sender cannot benefit from sequential persuasion, but in some cases with majority voting, he can. Finally, we analyzed aligned receivers with multi-dimensional states and have shown how to extend our Theorem 1 and why Theorem 2 (generically) does not work.

8 Appendix

8.1 Analysis of Binary Decisions

To understand better the conditions under which the simultaneous procedure achieves the first-best payoff, we consider a setup that is based on our examples. We assume that there are two receivers with the same threshold belief, α. They take actions $a_i = 1$ if their posterior belief is above α and $a_i = 0$ otherwise. The sender maximizes the sum of $a_i's$. We allow for imperfect correlation of the states. We assume without loss of generality that $x_1 \geq x_2$ and focus on the case where $x_2 \leq x_1 < \alpha$ (so the sender would like to persuade both receivers).

To describe the prior belief, we introduce the following notation. The overall state of the world is $\omega = (\omega_1, \omega_2)$. It can take the following values: $\omega_{ij} = (\omega_1 = i, \omega_2 = j)$ and the prior belief is $x_{ij} = Pr(\omega = \omega_{ij})$. For example, x_{HL} is the prior probability that $\omega_1 = H$ and $\omega_2 = L$.

In this setup, our result is that the simultaneous procedure fails to achieve the first-best outcome, if and only if $\alpha > \frac{1}{2}$ and ω_1 and ω_2 are sufficiently strongly negatively correlated. Formally we argue that:

Proposition 4. *The simultaneous procedure achieves the first-best outcome if and only if one of the following two inequalities hold:*

(i) $x_{HH} \geq x_{LH}\frac{2\alpha-1}{1-\alpha}$, *or*

(ii) $x_{HL} + x_{LH} \leq x_{LL}\frac{\alpha}{1-\alpha} + 2x_{HH}\frac{1-\alpha}{2\alpha-1}$.

While we provide formal proof below, we start with providing the main intuition. A simultaneous procedure is equivalent to choosing a joint distribution of signal S and state of the world ω. A simultaneous procedure achieves the first-best outcome if and only if the support of the marginal distribution of the posterior ω, conditional on $S = s_{ij}$, is $\{0, \alpha\}$.

Consider the following joint distribution of S and ω for some parameters $\{b, c, d, e, g\}$:

$S\backslash\omega$	ω_{HH}	ω_{HL}	ω_{LH}	ω_{LL}
s_{HH}	x_{HH}	b	b	c
s_{HL}	0	$x_{HL} - b$	0	d
s_{LH}	0	0	$x_{LH} - b$	e
s_{LL}	0	0	0	g

If a joint distribution that achieves the first-best payoff exists, then a signal S that has the distribution of this form (for appropriate values of the parameters) also achieves the first-best payoff. To see this, note that if it is possible to achieve the first-best payoff, it is sufficient to do it with a signal S that has at most four realizations that correspond to the posteriors that induce the receivers to take one of the four action combinations. Both receivers take action $a_1 = 1$ if the signal is s_{HH}. Only receiver 1 takes the desired action if the signal is s_{HL}. Only receiver 2 takes the desired action if the signal is s_{LH}. Finally, if the signal is s_{LL}, no receivers take the desired action. For the first-best payoff, it has to be that when both receivers take action $a_i = 0$, their posterior belief is that both states are L (which we denote ω_{LL}). This explains the necessity of the last row of this joint probability distribution. Similarly, when only one of the receivers takes $a_i = 0$ the belief that $\omega_i = H$ has to be α and the belief about the other state has to be 0, which explains the other zeros in the table.

Finally, consider the joint probability $Pr(s_{HH} \cap \omega_{HH})$ (the top left entry in the table). The reason it must be equal to x_{HH} is that when the state is ω_{HH} in the first-best outcome, both receivers take action $a_i = 1$ for sure. If that entry were not equal to x_{HH}, there would be a signal realization after which at least one of the receivers would not take the desired action, despite assigning a positive probability to both states being H. Such a procedure would fail to achieve the first best.

With this observation, the rest of the proof of Proposition 4 consists of identifying conditions in terms of α and the prior distribution of ω for which there exist parameters $\{b, c, d, e, g\}$ such that such a joint distribution is feasible and achieves the first-best outcome.

Formally we argue the following:

Proof of Proposition 4

Proof. 'If' direction.

Say condition (i) $x_{HH} \geq x_{LH}\frac{2\alpha-1}{1-\alpha}$ holds. Then the following signal is feasible: $b = x_{LH}$, $c = x_{HH}\frac{1-\alpha}{\alpha} + x_{LH}\frac{1-2\alpha}{\alpha}$, $d = [x_{HL} - x_{LH}]\frac{1-\alpha}{\alpha}$, $e = 0$ and $g = x_{LL} - c - d$.

First, we show that the suggested signal satisfies obedience constraints.[17]

$$x_{s_{HH},i} = \frac{x_{HH} + x_{LH}}{x_{HH} + x_{LH} + x_{LH} + x_{HH}\frac{1-\alpha}{\alpha} + x_{LH}\frac{1-2\alpha}{\alpha}} = \frac{x_{HH} + x_{LH}}{(x_{HH} + x_{LH})\frac{1}{\alpha}} = \alpha \tag{3}$$

$$x_{s_{HL},1} = \frac{x_{HL} - x_{LH}}{(x_{HL} - x_{LH}) + (x_{HL} - x_{LH})\frac{1-\alpha}{\alpha}} = \frac{x_{HL} - x_{LH}}{(x_{HL} - x_{LH})\frac{1}{\alpha}} = \alpha \tag{4}$$

Equation (3), for example, shows that given the signal realization s_{HH} expectation for dimension i is α, for $i = 1, 2$.

[17] Obedience constraints mean that the receivers find it optimal to take actions $(a_1, a_2) = (k, l)$ when the signal realization is s_{kl}.

Now we show that the suggested signal is also a feasible joint distribution. To do this we must show that $c + d + e \leq x_{LL}$. Expressing x_{LL} as a complementary probability, yields $x_{LL} = 1 - x_{HH} - x_{HL} - x_{LH}$.

Therefore, $c + d + e \leq x_{LL} \iff$

$$x_{HH}\frac{1-\alpha}{\alpha} + x_{LH}\frac{1-2\alpha}{\alpha} + [x_{HL} - x_{LH}]\frac{1-\alpha}{\alpha} \leq 1 - x_{HH} - x_{HL} - x_{LH}. \tag{5}$$

This expression simplifies to:

$$x_{HH}\frac{1}{\alpha} + x_{HL}\frac{1}{\alpha} \leq 1, \tag{6}$$

which holds, since we assume that $x_i \leq \alpha$.

Now suppose that (i) is violated, but condition (ii): $x_{HL} + x_{LH} \leq x_{LL}\frac{\alpha}{1-\alpha} + 2x_{HH}\frac{1-\alpha}{2\alpha-1}$ is satisfied. Then the following signal is feasible: $b = x_{HH}\frac{1-\alpha}{2\alpha-1}$, $c = 0$, $d = (x_{HL} - x_{HH}\frac{1-\alpha}{2\alpha-1})(\frac{1-\alpha}{\alpha})$ and $e = (x_{LH} - x_{HH}\frac{1-\alpha}{2\alpha-1})(\frac{1-\alpha}{\alpha})$.

First, we show that obedience constraints are satisfied.

$$x_{s_{HH},i} = \frac{x_{HH} + x_{HH}\frac{1-\alpha}{2\alpha-1}}{x_{HH} + x_{HH}\frac{1-\alpha}{2\alpha-1} + x_{HH}\frac{1-\alpha}{2\alpha-1}} = \alpha \tag{7}$$

$$x_{s_{HL},1} = \frac{x_{HL} - x_{HH}\frac{1-\alpha}{2\alpha-1}}{x_{HL} - x_{HH}\frac{1-\alpha}{2\alpha-1} + (x_{HL} - x_{HH}\frac{1-\alpha}{2\alpha-1})\frac{1-\alpha}{\alpha}} = \frac{1}{1+\frac{1-\alpha}{\alpha}} = \alpha \tag{8}$$

$$x_{s_{LH},2} = \frac{x_{LH} - x_{HH}\frac{1-\alpha}{2\alpha-1}}{x_{LH} - x_{HH}\frac{1-\alpha}{2\alpha-1} + (x_{LH} - x_{HH}\frac{1-\alpha}{2\alpha-1})\frac{1-\alpha}{\alpha}} = \frac{1}{1+\frac{1-\alpha}{\alpha}} = \alpha. \tag{9}$$

Feasibility means that $d + e \leq x_{LL}$, i.e.:

$$(x_{HL} - x_{HH}\frac{1-\alpha}{2\alpha-1})(\frac{1-\alpha}{\alpha}) + (x_{LH} - x_{HH}\frac{1-\alpha}{2\alpha-1})(\frac{1-\alpha}{\alpha}) \leq x_{LL} \tag{10}$$

which is equivalent to condition (ii):

$$x_{HL} + x_{LH} \leq x_{LL}\frac{\alpha}{1-\alpha} + 2x_{HH}\frac{1-\alpha}{2\alpha-1}. \tag{11}$$

'Only if' direction.

If both conditions (i) and (ii) do not hold, then we claim that, for any feasible signal (S), at least one of the obedience constraints does not bind. This is so because if all constraints bind, then $c + d + e > x_{LL}$, i.e., the signal (S) is not a feasible joint distribution. We now prove this claim.

Since condition (i) is violated, we consider the following recommendation rule:

$b = \tilde{b} := x_{HH}\frac{1-\alpha}{2\alpha-1}$, $c = \tilde{c} := 0$, $d = \tilde{d} := (x_{HL} - x_{HH}\frac{1-\alpha}{2\alpha-1})(\frac{1-\alpha}{\alpha})$ and $e = \tilde{e} := (x_{LH} - x_{HH}\frac{1-\alpha}{2\alpha-1})(\frac{1-\alpha}{\alpha})$.

We have shown in equations (7) − (9), that the suggested signal satisfies the obedience constraints. The feasibility constraint becomes: $\tilde{c} + \tilde{d} + \tilde{e} \leq x_{LL}$, which upon substituting for $\tilde{c} + \tilde{d} + \tilde{e}$ yields

$$(x_{HL} - x_{HH}\frac{1-\alpha}{2\alpha-1})(\frac{1-\alpha}{\alpha}) + (x_{LH} - x_{HH}\frac{1-\alpha}{2\alpha-1})(\frac{1-\alpha}{\alpha}) \leq x_{LL}. \tag{12}$$

(12) is equivalent to condition (ii), which is a contradiction.

Now we show that no other feasible signal exists that achieves the first-best payoff. Thus, we consider a signal with $c = c' > 0$. Then, the following is true:

$$c' + d' + e' > \tilde{c} + \tilde{d} + \tilde{e} > x_{LL}, \tag{13}$$

where c', d', and e' are the values of the suggested new signal. Condition (13) holds because for any signal that achieves the first-best payoff, b is decreasing in c, whereas d and e are decreasing in b. To see this, suppose that $c = c' > \tilde{c}$, then we claim that $b' < \tilde{b}$. If not, then

$$x_{s_{HH},i} = \frac{x_{HH} + x_{HH}\frac{1-\alpha}{2\alpha-1}}{x_{HH} + x_{HH}\frac{1-\alpha}{2\alpha-1} + x_{HH}\frac{1-\alpha}{2\alpha-1} + c'} < \alpha, \tag{14}$$

and the obedience constraint for s_{HH} is violated. If a recommendation rule achieves the first-best payoff, then $d' = (x_{HL} - b')(\frac{1-\alpha}{\alpha})$ and $e' = (x_{LH} - b')(\frac{1-\alpha}{\alpha})$, i.e., the obedience constraints for s_{HL} and s_{LH} bind, so $b' < \tilde{b}$ implies that $d' > \tilde{d}$ and $e' > \tilde{e}$, as we claimed, establishing (13). This completes the argument.

This means that if both conditions are violated, then there does not exist a feasible signal that induces a distribution of marginal posteriors with support $\{0, \alpha\}$. □

8.2 Additional Proofs.

Proof that the sequential procedure in Example 1 is optimal.

Proof. As we discussed in Section 5.1, when the states are perfectly correlated, the highest payoff from sequential persuasion is

$$U^*_{Seq}(x) = max\{cav(f_1 + cav f_2)(x), cav(f_2 + cav f_1)(x)\},$$

Since in this problem the two actions are symmetric, we need only consider a single order.

We have

$$\begin{aligned}(f_1 + cav f_2)(x) &= 1 + \frac{x}{0.8} \text{ if } x \leq 0.2\\ &= \min\{\frac{x}{0.8}, 1\} \text{ if } x > 0.2.\end{aligned}$$

Concavification of that function gives us:

$$\begin{aligned}cav(f_1 + cav f_2)(x) &= 1 + \frac{x}{0.8} \text{ if } x \leq 0.2\\ &= 1 + \frac{1}{4}\frac{1-x}{0.8} \text{ if } x > 0.2\end{aligned}$$

In particular, if we start with a prior $x = \frac{1}{2}$ the highest sequential payoff is

$$U^*_{Seq}(p) = 1 + \frac{1}{4}\frac{1-0.5}{0.8} = 1.1563.$$

□

Lemma 3. *Suppose that X is a random variable and there are two values $x_1 < x_2 \in \mathbb{R}$ such that:*

- $Pr(X \in (x_1, x_2)) = 0$.
- $E[X] \in (x_1, x_2)$

Then there exists a binary random variable X' which is a garbling of X with realizations x'_1, x'_2, such that $E[X|X' = x'_1] = x_1$, and $E[X|X' = x'_2] = x_2$.

Proof. We first note that $Pr(X \leq x_1), Pr(X \geq x_2) > 0$. We also argue that without loss of generality, we can consider binary signals S with only two realizations a and b where $a \leq x_1 < x_2 \leq b$. This follows from the fact that we can define a signal $\hat{S}$ where $\hat{S} = E(S|S \leq x_1)$ when $S \leq x_1$ and $\hat{S} = E(S|S \geq x_2)$ where $S \geq x_2$. Since $\hat{S}$ is garbling of S, garbling of $\hat{S}$ is also garbling of S.

Consider a signal S_z indexed by z; the signal is binary with realizations $\{s_1, s_2\}$ that occur with probabilities $\{1-z, z\}$, respectively. Specifically, s_2 occurs with probability $\alpha \cdot z$ when $S = a$ and with probability $(1-\alpha) \cdot z$ when $S = b$, where $\alpha \in [0, 1]$ is defined by:

$$x_2 = (1-\alpha) \cdot a + \alpha \cdot b.$$

As a result we have that $\forall z$: $E(S|S_z = s_2) = x_2$. The feasible range for z is $[0, min\{\frac{Pr(S=x_2)}{\alpha}, \frac{Pr(S=x_1)}{1-\alpha}\}]$. We first note that $\frac{Pr(S=x_2)}{\alpha} \leq \frac{Pr(S=x_1)}{1-\alpha}$. This implies that the feasible range for z is $[0, \frac{Pr(S=x_2)}{\alpha}]$.

To see why this holds, note that:

$$\frac{Pr(S=x_2)}{\alpha} > \frac{Pr(S=x_1)}{1-\alpha} \Rightarrow \frac{Pr(S=x_2)}{Pr(S=x_1)} > \frac{\alpha}{1-\alpha}.$$

This would imply that $E(S) > x_2$ which is a contradiction. Consider the other realization s_1 and let $H(z) \equiv E(S|S_z = s_1)$. The claim follows from the intermediate value theorem as $H(0) = E(S)$ and $H(\frac{Pr(S=x_2)}{\alpha}) = a < x_1$. Hence, we conclude that there exists z^* so that $H(z^*) = E(S|S_{z^*} = s_1) = x_1$. □

Proof of Lemma 2

Proof. Note that:

$$\begin{aligned} h_{i,j}(p) &= p \cdot Pr(\omega_j = H|\omega_i = H) + (1-p) \cdot Pr(\omega_j = H|\omega_i = L) \\ &= Pr(\omega_j = H|\omega_i = L) + p \cdot [Pr(\omega_j = H|\omega_i = H) - Pr(\omega_j = H|\omega_i = L)] \end{aligned}$$

The proof then follows from $Pr(\omega_j = H|\omega_i = H), Pr(\omega_j = H|\omega_i = L) \in [0,1]$. □

Proof of Proposition 3

Proof. Part 1) First, suppose the simultaneous procedure induces (with positive probability) a posterior belief such that some receivers accept and some do not. Then a sequential procedure can further improve the sender's payoff by letting those that accept in the original case accept first and then reveal the ω, causing some additional receivers to accept with positive probability.[18]

Second, suppose that with a positive probability, the simultaneous procedure induces a posterior belief such that all receivers accept. Then, note that the line that connects that belief and the prior belief has to cross the acceptance boundaries of all receivers (with the possible exception of the receiver with the smallest acceptance set). Then the following sequential procedure improves upon the simultaneous procedure: Start with garbling all the signals so that instead of the posterior being in the smallest acceptance set, it would be on the boundary of the receiver with the largest acceptance set. If that belief is realized, the receiver with the largest acceptance set is asked to accept first, and then the sender reveals the original signal. That increases the probability of acceptance of one receiver without changing the probability of acceptance of all the other receivers.

Part 2) Suppose without loss of generality that the optimal procedure does not induce in the first round beliefs on the boundary of the receiver with the largest acceptance set.

[18] This step uses the assumption that all acceptance sets contain the same set of extreme beliefs, so if not all receivers accept, the beliefs assign positive but less than one probability to the acceptance states.

Then using the same garbling argument as in part 1, the sender can improve its payoff by increasing the probability of acceptance by that receiver. By inductive argument, the rest of order follows.

Part 3) In the first best, where the sender persuades each of the receivers privately, the optimal persuasion of each receiver induces binary beliefs: One on the (closure of the) fully mixed boundary of the acceptance set and one outside the acceptance set. The line connecting these two beliefs crosses through the prior. Generically, these lines do not coincide for different receivers. Hence, even a fully flexible procedure would not be able to replicate the outcome of the first-best procedure: To replicate the first-best for receiver i, it would need to induce a particular belief on the boundary of that receiver's acceptance set. But then, generically, this belief would not be on the correct line connecting the prior and the optimal point on the boundary of receiver j. □

References

Alonso, R. and O. Câmara (2016). Persuading voters. *American Economic Review 106*(11), 3590–3605.

Anthony-North, V. (2022). Doubling down: Inconsistent prosecutions, capital punishment, and double jeopardy. *NYUL Rev. 97*, 235–274.

Arieli, I. and Y. Babichenko (2019). Private bayesian persuasion. *Journal of Economic Theory 182*(1), 185–217.

Aumann, R. J., M. Maschler, and R. E. Stearns (1995). *Repeated games with incomplete information*. MIT press.

Bergemann, D. and S. Morris (2016a). Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics 11*(2), 487–522.

Bergemann, D. and S. Morris (2016b). Information design, bayesian persuasion, and bayes correlated equilibrium. *American Economic Review 106*(5), 586–91.

Bergemann, D. and S. Morris (2019). Information design: a unified perspective. *Journal of Economic Literature 57*, 1–57.

Board, S. and J. Lu (2018). Competitive information disclosure in search markets. *Journal of Political Economy 126*(5), 1965–2010.

Ely, J. C. (2017). Beeps. *American Economic Review 107*(1), 31–53.

Ely, J. C. and M. Szydlowski (2020). Moving the goalposts. *Journal of Political Economy 128*(2), 468–506.

Gamp, T. (2019). Guided search. *Working paper*.

Hörner, J. and A. Skrzypacz (2016). Selling information. *Journal of Political Economy 124*(6), 1515–1562.

Kamenica, E. (2019). Bayesian persuasion and information design. *Annual Review of Economics 11*, 249–272.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *American Economic Review 101*(6), 2590–2615.

Li, F. and P. Norman (2021). Sequential persuasion. *Theoretical Economics 16*, 639–675.

Orlov, D., A. Skrzypacz, and P. Zryumov (2020). Persuading the principal to wait. *Journal of Political Economy 128*(7), 2542–2578.

Rayo, L. and I. Segal (2010). Optimal information disclosure. *Journal of Political Economy 118*(5), 949–987.

Smolin, A. (2021). Dynamic evaluation design. *American Economic Journal: Microeconomics 13*, 300–331.

Turlo, S. (2019). Persuasion through ordered information. *Working paper*.