Theory and Decision
https://doi.org/10.1007/s11238-025-10046-y

Persuasion with simplistic worldviews

Maxim Senkov[1] · **Toygar T. Kerman**[2]

Accepted: 26 April 2025

Abstract
We study a Bayesian persuasion model with two-dimensional states of the world, in which the sender (she) and receiver (he) have heterogeneous prior beliefs and care about different dimensions. The receiver is a naive agent who has a *simplistic worldview*: while he knows the correct marginal distributions of the dimensions, he ignores the dependency between the two dimensions of the state. As a benchmark, we first characterize optimal disclosure when the receiver is rational, i.e. when the receiver knows the correct prior distribution and thus shares a common prior with the sender. Then, we provide a characterization for the sender's gain from persuasion when the receiver is naive, and contrast our results to the benchmark. Finally, we show that the receiver benefits from having a simplistic worldview if and only if it makes him perceive the states in which his interest is aligned with the sender as less likely.

Keywords Bayesian persuasion · Correlation neglect · Information design · Misspecified prior

1 Introduction

Standard models of Bayesian persuasion assume that the sender (she) and receiver(s) (he) care about the same dimension of the true state of the world. For example, in their seminal paper Kamenica & Gentzkow (2011) consider a prosecutor (sender) and a judge (receiver), both of whom are only interested in one dimension of the

The author acknowledges funding by the Hungarian National Research, Development and Innovation Office, Project Number K-143276.

✉ Toygar T. Kerman
toygar.kerman@uni-corvinus.hu

Maxim Senkov
maxim.senkov@eruni.org

[1] European Research University, Sokolská třída 2416/33, 702 00 Ostrava, Czech Republic

[2] Institute of Economics, Corvinus University of Budapest, Fővám tér 8, 1093 Budapest, Hungary

Published online: 22 May 2025

true state, that is whether the defendant is guilty or innocent. What if the sender and receiver are concerned about different dimensions of the true state?

Consider a company that wishes to build a new factory and only cares about the project's profitability. In order to be perceived as environmentally friendly, the company asks a sustainability consultant to evaluate the project and provide a recommendation on whether to continue or terminate it. However, the consultant only cares about a different aspect of the project: its sustainability. Then how should the consultant conduct her research to convince the company to continue the project only when it is sustainable?

When considering multidimensional states of the world, however, a behavioral aspect emerges: thinking about how multiple issues relate to each other is cognitively demanding. Experimental and psychological evidence suggest that people find it challenging to work with joint distributions of random variables (Shimojo & Ichikawa, 1989; Zhao, Shah, and Osherson, 2009; Wolfe & Reyna, 2010) and tend to underestimate or ignore the correlation between state variables or observed signals (Kallir & Sonsino, 2009; Eyster & Weizsäcker, 2011; Enke & Zimmermann, 2019).

In this paper, we study how the receiver's *simplistic worldview* (i.e. treating dependent state variables as independent) affects the sender's ability to influence his decision. We extend Kamenica & Gentzkow (2011) by (*i*) allowing for two-dimensional states of the world and (*ii*) considering a behavioral receiver who ignores the dependency of state variables. In particular, we model (*ii*) by assuming the receiver has a misspecified prior belief; while the receiver knows the correct marginal distribution for each dimension, he assumes that they are independent (which leads to heterogeneous priors). We first characterize optimal disclosure when the receiver is *rational* (i.e. he does not ignore dependency) as a benchmark. Further, we characterize the optimal signal for the case of a naive receiver (i.e. he assumes that dimensions are independent), as well as for a receiver who has an arbitrary prior belief. Finally, we analyze the welfare effects of the receiver's simplistic worldview and provide comparative statics on the receiver's prior belief.

1.1 Illustrative example

Let us return to our example of a company and sustainability consultant. Both the company and the consultant are initially uncertain whether the project is profitable (P) or loss-making (L) and sustainable (S) or unsustainable (U). The company chooses to either continue or terminate the project and only cares about its profitability. Continuing the project when it is profitable and terminating it when it is loss-making provide the company a utility of 1, otherwise, the company's utility is 0. On the other hand, the consultant only cares about the sustainability of the project. Continuing the project when it is sustainable and terminating it when it is unsustainable provide the consultant a utility of 1, otherwise the consultant's utility is 0.

Hence, the true state of the world has two dimensions, and the conflict of interest depends on the joint distribution of the dimensions. First, suppose that the company is rational. That is, it shares a common prior with the consultant, where the

Table 1 Prior belief distribution

	L	P	
U	0.5	0.15	0.65
S	0.2	0.15	0.35
	0.7	0.3	

prior distribution of the states is given as follows (with marginal distributions on the sides) (Table 1).

The company asks the consultant to investigate the profitability and sustainability of the project and provide a recommendation to either *continue* (c) or *terminate* (t) it. Suppose that if the company is indifferent between continuing and terminating the project, it continues. Since the project is initially more likely to be loss-making, the company would (ex ante) prefer the project to be terminated (which yields expected payoff 0.65 to the consultant and 0.7 to the company). However, the consultant can improve her expected payoff by strategically disclosing information about the true state of the world. Mathematically, the information disclosure can be represented by a *signal* π that sends recommendations contingent on the state of the world. Consider an optimal signal given as follows.

π	(U, L)	(U, P)	(S, L)	(S, P)
c	0	0.33	1	1
t	1	0.67	0	0

First, observe that π recommends continuing with probability 1 when the project is both sustainable and profitable (i.e. in state (S, P)) and recommends terminating with probability 1 when the project is both unsustainable and loss-making (i.e. in state (U, L)). This is intuitive since in these states the preferred actions of the sender and receiver are aligned. However, observe that π recommends to continue the project with probability 1 also in state (S, L).[1] For this recommendation to be persuasive, in state (U, P) the sender recommends to continue with a relatively high probability of 0.33, despite the fact that she prefers to terminate since the project is unsustainable. In particular, the consultant chooses the probability of c in state (U, P) such that the company's posterior belief that the project is profitable is exactly 1/2 and thus it chooses to continue. While the consultant's expected payoff of employing π is 0.95, which improves upon the case of no disclosure, the company's expected payoff from π is 0.7, which is equal to the case of no disclosure.

Now suppose that the company is naive and ignores the dependency between the states of the world due to insufficient knowledge of sustainability. In other words, the company assumes that the dimensions of the true state are independent, and thus has a misspecified prior. While the company

[1] Note that there also exists an optimal signal which does not recommend to continue with probability 1 in state (S, L).

Table 2 Misspecified prior belief distribution

	L	P	
U	0.455	0.195	0.65
S	0.245	0.105	0.35
	0.7	0.3	

knows the marginal distributions $\Pr(S) = 0.35$ and $\Pr(P) = 0.3$, it assumes that $\Pr(S, P) = \Pr(S)\Pr(P) = (0.35) \cdot (0.3) = 0.105$.[2] The company's *simplistic worldview* (i.e. misspecified prior) is given as follows (Table 2).

Notice that the company now perceives states (S, L) and (U, P) as relatively more likely than before. Therefore, the sender has two options to keep the recommendation persuasive: (*i*) recommend to terminate more often than before when the project is loss-making (i.e. decrease probability of c in (S, L)) or (*ii*) recommend to continue more often than before when the project is profitable (i.e. increase probability of c in (U, P)). In other words, the consultant needs to make the recommendation to continue *more informative* regarding the project's profitability. It turns out that given the misspecified prior of the receiver, it is optimal for the sender to follow the second route.

$\hat{\pi}$	(U, L)	(U, P)	(S, L)	(S, P)
c	0	0.72	1	1
t	1	0.28	0	0

Indeed, when the project is unsustainable and profitable, the recommendation to continue is sent with probability 0.72, while it was sent with probability 0.33 in this state before. The consultant's expected payoff of employing $\hat{\pi}$ is 0.892, still improving upon the case of no disclosure, albeit less than when the receiver is rational. On the other hand, the company's expected payoff from $\hat{\pi}$ is 0.758, which is higher relative to the case of having a correct prior belief (and also higher relative to the case of no disclosure).[3] Hence, while the company's simplistic worldview *benefits* the company, it is *detrimental* to the consultant.

The example illustrates how the receiver *benefits* from ignoring dependency between the state variables. However, both the sender's approach regarding disclosure in states with misaligned preferences and whether the receiver benefits from ignoring dependency crucially depend on the prior belief distribution. In fact, the receiver might also be *worse off* due to ignoring the dependency of states. In our analysis, we consider all possible prior belief distributions, provide a full

[2] That is, the company assumes that $\Pr(P) = \Pr(P|S)$.

[3] Here, we use a misspecified model framework, which implies the existence of a correct prior belief (corresponding to the sender's belief). Thus, the naive receiver's expected payoff is evaluated using the correct prior belief.

characterization of the sender's gain from persuasion, and pin down the conditions under which the receiver benefits from having a simplistic worldview.

1.2 Related literature

Our paper extends Kamenica & Gentzkow (2011) in two aspects and contributes to three strands of literature. First, it relates to the literature on misspecified models. The receiver in our model knows the correct marginal distribution of each dimension, yet, assumes that the dimensions are independent.[4] Hence, the receiver possesses a misspecified prior belief about the states of the world. This naturally relates to studies that consider correlation neglect, as they motivate the behavioral assumption we make about the receiver. One main difference from such models (e.g. Levy et al. 2022) is that the receiver in our model ignores dependency between *states*, whereas most papers in the literature focus on the receiver neglecting correlation between *messages*.[5]

Our findings share similarities with those from studies focusing on the political applications of correlation neglect. Ortoleva and Snowberg (2015) show that correlation neglect might lead to overconfidence in voters and that overconfidence is an important predictor for ideological extremeness. In our model, the misspecified prior belief of the receiver leads to the receiver having over-optimistic or over-pessimistic beliefs about the state of the world. Levy and Razin (2015) show that voters neglecting the correlation between their information sources might contribute to higher political polarization in society, howbeit, voters might benefit from it. Similarly, we illustrate that if the receiver ignores dependency between states then the optimal signal might be more informative.[6]

Our behavioral assumption regarding the receiver alters his perception of the joint belief distribution. Yet, there is no ambiguity about the message the receiver observes. Eliaz et al. (2021a) consider a sender who can redact his message once it is chosen, but the receiver only knows the message strategy and not the redaction strategy. In a related study, Eliaz et al. (2021b) consider a sender who can attach interpretations to messages. Unlike our model where the receiver's perception of the signal structure remains correct, both studies assume a distortion in the perception of the signal structure.

[4] The study of errors in estimating the conjunctive probability of events has been prominent also in the cognitive science literature (Wyer Jr, 1976; Tversky & Kahneman, 1983; Khemlani, Lotstein, and Johnson-Laird, 2015).

[5] Eyster and Rabin (2005) introduce *cursed equilibrium* which assumes that players underestimate the correlation between other players' actions and information.

[6] Che and Kartik (2009) study a disclosure game between a sender acquiring costly information and a receiver, and also find that heterogeneity in the priors of the sender and receiver leads to more information being disclosed. This also relates to studies that investigate: the informativeness of communication protocols in different environments (Gentzkow & Kamenica, 2016; Ichihashi, 2019; Brooks, Frankel, and Kamenica, 2022; Herings, Karos, and Kerman, 2024), how a receiver can strategize to increase the informativeness of the signal (Austen-Smith & Banks, 2000; Kartik, 2007; Ambrus & Egorov, 2017; Tsakas, Tsakas, and Xefteris, 2021), and how different kinds of biases or reputation concerns might affect persuasion (Hagmann & Loewenstein, 2017; Augias & Barreto, 2020; de Clippel & Zhang, 2022; Li, 2022).

Second, our paper contributes to the literature on Bayesian persuasion with heterogeneous priors. The study that comes closest to ours is Alonso and Câmara (2016), which consider a general setup and mainly focus on conditions under which persuasion has value for the sender. Thus, their insights are particularly illuminating in the context of state-independent (sender) preferences. Meanwhile, we focus on a sender with state-dependent preferences, so our sender always benefits from persuasion. Moreover, we study the implications of interplay between the ex ante preference alignment (as captured by the prior) and the heterogeneous priors for the sender's persuasion capability. Another difference with our study is that they do not consider an interpretation of a finite state space as two correlated random variables, and thus do not focus on the effect of correlation neglect. Meanwhile, this constitutes one of the main focus points of our paper. Finally, the specific assumptions we make allow us to characterize the specific form of optimal signals for arbitrary prior beliefs.

Kosterina (2022) analyzes optimal disclosure when there is ambiguity about the receiver's prior. In particular, the sender only knows that the receiver assigns probabilities to the states that are above a certain threshold.[7] While we also consider a sender and receiver who disagree on the prior, in our setup this is also common knowledge. In a result similar to ours, the author shows that the effect of ignorance on the receiver's welfare is ambiguous. However, while in her model the receiver always strictly prefers an ignorant sender over a perfectly informed one, in our model the sender *might* prefer a receiver that has a correct prior to one that does not.

Third, our study relates to the literature on persuasion with multidimensional state variables. Rayo and Segal (2010) consider a model of information disclosure, in which the sender's expected payoff is the product of posterior mean of the sender's and receiver's dimension of state, and characterize optimal disclosure. Tamura (2018) studies optimal disclosure in a more general setup similar to the model in Rayo and Segal (2010). In a related study, Dworczak and Kolotilin (2019) characterize the optimal persuasion mechanism when the state is two-dimensional and the objective is quadratic. Malamud and Schrimpf (2021) study a model where the sender observes the multidimensional state and show that it is optimal for the sender to "reduce" the dimensions of the state. All studies above consider a rational sender and receiver. In contrast, our paper considers implications of the receiver's ignorance regarding the connection between the dimensions of state for the optimal signal and welfare. Babichenko et al. (2022) consider a sender who provides information to a receiver about a product, where each dimension represents a different attribute of the product. The authors extend the standard model also by assuming the sender does not know the utility of the receiver and wishes to minimize regret and show that ignorance of the utility of the receiver is extremely harmful to the sender. We, on the other hand, assume that agents have complete information about the utility functions, yet, the receiver incorrectly assumes independence of the dimensions.

[7] In a related study, Galperti (2019) considers a setup in which the receiver assigns very low (or zero) probability to some states of the world and the receiver is persuaded in those states only when the evidence is sufficiently surprising.

2 Preliminaries

2.1 The model

A sender (she) wishes to persuade a receiver (he) to take an action depending on the true state of the world. Let $\Omega = \Sigma \times R$ be the set of *states* of the world, such that the sender's payoff depends only on $\Sigma = \{\sigma_0, \sigma_1\}$ and the receiver's payoff depends only on $R = \{\rho_0, \rho_1\}$. For each $k, \ell \in \{0, 1\}$, let $\omega_{k\ell} \equiv (\sigma_k, \rho_\ell) \in \Omega$.

We start by considering a rational receiver, hence, the sender and receiver share a common prior belief $\mu \in \Delta(\Omega)$ about the true state of the world. For each $k, \ell \in \{0, 1\}$, let $\mu_{k\ell} \equiv \Pr(\sigma_k, \rho_\ell)$ and let the marginal distributions be denoted as $\mu(\sigma_k) \equiv \Pr(\sigma_k)$ and $\mu(\sigma_\ell) \equiv \Pr(\rho_\ell)$. The distribution μ is presented in Table 3 with marginal distributions given on the sides.

Communication, actions, and beliefs. The receiver's action set is binary and given by $A = \{0, 1\}$. A signal consists of a set of messages $M = \{0, 1\}$ and a family of distributions $\{\pi(\cdot|\omega)\}_{\omega\in\Omega}$ over M. That is, $\pi : \Omega \to \Delta(M)$ maps each state of the world to a probability distribution over the two messages. Thus, we denote a signal by $\pi = (\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$ where $\pi_{k\ell} \equiv \pi(1|\omega_{k\ell})$, for $k, \ell \in \{0, 1\}$ and the set of all signals by Π.

Given $\mu \in \Delta(\Omega)$ and $\pi \in \Pi$, a message $m \in M$ generates the posterior belief $\mu^m \in \Delta(\Omega)$. For each $k, \ell \in \{0, 1\}$, the marginal posterior beliefs upon observing message m are given by $\mu^m(\sigma_k)$ and $\mu^m(\rho_\ell)$, where

$$\mu^1(\rho_\ell) = \frac{\mu_{k\ell}\pi_{k\ell} + \mu_{\ell\ell}\pi_{\ell\ell}}{\mu_{k\ell}\pi_{k\ell} + \mu_{\ell\ell}\pi_{\ell\ell} + \mu_{\ell k}\pi_{\ell k} + \mu_{kk}\pi_{kk}},$$

$$\mu^0(\rho_\ell) = \frac{\mu_{k\ell}(1-\pi_{k\ell}) + \mu_{\ell\ell}(1-\pi_{\ell\ell})}{\mu_{k\ell}(1-\pi_{k\ell}) + \mu_{\ell\ell}(1-\pi_{\ell\ell}) + \mu_{\ell k}(1-\pi_{\ell k}) + \mu_{kk}(1-\pi_{kk})},$$

and $\mu^m(\sigma_k)$ is given analogously.[8] We denote the posterior belief for the joint distribution upon observing message m by $\mu^m(\sigma_k, \rho_\ell)$.

Payoffs. The receiver wishes his action to match the dimension of the state he cares about, i.e., the receiver wants to choose 0 when the state is (σ_k, ρ_0) and wants to choose 1 when the state is (σ_k, ρ_1), for any $k \in \{0, 1\}$. The receiver's utility function $u : A \times R \to \{0, 1\}$ is given by

$$u(a, \rho) = \begin{cases} 1 \text{ if } (a, \rho) = (0, \rho_0) \text{ or } (a, \rho) = (1, \rho_1), \\ 0 \text{ otherwise.} \end{cases}$$

Given the receiver's preferences, it follows that the receiver is indifferent between 0 and 1 if $\mu^m(\rho_0) = \mu^m(\rho_1) = 1/2$.

The sender wishes the chosen action to match the dimension of the state she cares about. The sender's utility function $v : A \times \Sigma \to \{0, 1\}$ is given by

[8] We drop π when denoting $\mu^m(\sigma_k)$ and $\mu^m(\rho_\ell)$ to ease notation.

Table 3 Common prior μ

(Σ, R)	ρ_0	ρ_1	
σ_0	μ_{00}	μ_{01}	$\mu(\sigma_0)$
σ_1	μ_{10}	μ_{11}	
	$\mu(\rho_0)$		

$$v(a,\sigma) = \begin{cases} 1 \text{ if } (a,\sigma) = (0,\sigma_0) \text{ or } (a,\sigma) = (1,\sigma_1), \\ 0 \text{ otherwise.} \end{cases}$$

The following assumption on the prior beliefs implies that the default action of the receiver is 0, which is without loss of generality.

Assumption 1 $\mu(\rho_1) < \mu(\rho_0)$.

Given $\pi \in \Pi$, the *choice rule* $\alpha^\pi : M \to A$ of the receiver is given by

$$\alpha^\pi(m) = \begin{cases} 1 \text{ if } \mu^m(\rho_1) \geq \frac{1}{2}, \\ 0 \text{ otherwise.} \end{cases} \tag{1}$$

That is, the receiver switches his action if he receives a message that leads to a posterior of at least 1/2 about state being ρ_1.

Note that considering a binary message space $M = \{0, 1\}$ is without loss of generality, since the receiver's action set is binary as well. In the case where the sender and the receiver share a prior belief (i.e. the receiver is rational) it follows from a simple argument analogous to the revelation principle, as in Kamenica and Gentzkow (2011), that the sender can restrict attention to straightforward signals without loss of generality.[9] On the other hand, if there is disagreement about the prior (i.e. the receiver is naive) the optimality of straightforward signals follows from the message space reduction argument of Proposition 2 (i) in Alonso and Câmara (2016). Hence, messages 1 and 0 are interpreted as recommendations to take action 1 and 0, respectively.

2.2 Benchmark

We start by characterizing optimal disclosure when the receiver does not ignore the dependency of states, i.e., when he is rational. This will serve as a benchmark for the next sections.

Depending on the prior, the sender may choose $\pi \in \Pi$ to induce a switch of the receiver's default action 0 to 1. For this to be successful, message 1 should induce a posterior belief $\mu^1(\rho_1) \geq 1/2$, i.e. the receiver should follow the recommendation.

[9] A signal is *straightforward* if it holds that (*i*) $M \subseteq A$ and (*ii*) for all $a \in M$, $\alpha^\pi(a) = a$.

Lemma 1 *(Straightforwardness constraint)* Given $\pi \in \Pi$, the receiver chooses action 1 upon observing message 1 if and only if

$$\mu_{00}\pi_{00} + \mu_{10}\pi_{10} \leq \mu_{01}\pi_{01} + \mu_{11}\pi_{11}.$$

Note that Lemma 1 implies that the straightforwardness constraint for action 0 given by $\mu_{01}(1-\pi_{01}) + \mu_{11}(1-\pi_{11}) < \mu_{00}(1-\pi_{00}) + \mu_{10}(1-\pi_{10})$ is also satisfied. See Lemma A1 in the appendix for the formal statement and proof.

We can now state the sender's optimization problem, which is to maximize the ex ante expected utility by choosing a signal:

$$\begin{aligned} \max_{\pi \in \Pi} \ & \mu_{11}\pi_{11} + \mu_{10}\pi_{10} - \mu_{01}\pi_{01} - \mu_{00}\pi_{00} \\ & \text{s.t. } \mu_{00}\pi_{00} + \mu_{10}\pi_{10} \leq \mu_{01}\pi_{01} + \mu_{11}\pi_{11}. \end{aligned} \tag{2}$$

Let π^* denote an optimal signal when the receiver is rational.

When the straightforwardness constraint for message 1 is binding, the receiver is indifferent between choosing action 0 and action 1. However, recall that in this case the tie is broken in favor of action 1 (i.e. the receiver chooses 1 when $\mu^1(\rho_1) = 1/2$). The next lemma shows that this tie-breaking is sender-preferred. That is, the sender prefers action 1 when the realization of the optimal signal is 1.[10]

Lemma 2 Given $\pi^* \in \Pi$ that solves (2), it holds that $\mu^1(\sigma_1) \geq 1/2$.

In the appendix, we postpone the proof of Lemma 2 to after the proof of Proposition 1 as it employs the properties of an optimal signal.

We are now ready to present the characterization of optimal disclosure in the case of a rational receiver.

Proposition 1 Assume that the sender and receiver share a common prior μ. An optimal signal is given by

$$(\pi^*_{00}, \pi^*_{01}, \pi^*_{10}, \pi^*_{11}) = \begin{cases} (0,0,1,1) & \text{if } \mu_{11} \geq \mu_{10}, \\ (0,\beta,\gamma,1) & \text{otherwise,} \end{cases}$$

where $\mu_{10}\gamma - \mu_{01}\beta = \mu_{11}$.

There are a few aspects to note about Proposition 1. First, it is intuitive that $\pi^*_{00} = 0$ and $\pi^*_{11} = 1$, since in states (σ_0, ρ_0) and (σ_1, ρ_1) the sender's and receiver's preferences are aligned. In other words, it is optimal for the sender to *disclose the truth* in these states, i.e., she sends message 1 with probability 0 in state (σ_0, ρ_0) and with probability 1 in state (σ_1, ρ_1). More precisely, when the state is (σ_1, ρ_1) the sender has an incentive to *increase* the frequency of message 1, which switches the receiver's action. Similarly, when the state is (σ_0, ρ_0) the sender has an incentive to *decrease* the frequency

[10] Note that we do not need to consider the sender-optimal tie-breaking for message 0 because as Lemma A1 suggests, the receiver strictly prefers action 0 upon observing message 0.

of message 1. Note that both increasing π_{11} and decreasing π_{00} *relaxes* the straightforwardness constraint in (2).

Further, observe that while the sender has an incentive to *decrease* π_{01} in state (σ_0, ρ_1) and *increase* π_{10} in state (σ_1, ρ_0), this *tightens* the straightforwardness constraint since the sender's and receiver's preferences are misaligned. The sender's choice of π_{01} and π_{10} makes the constraint bind if $\mu_{11} \leq \mu_{10}$. Finally, if $\mu_{11} > \mu_{10}$, then the constraint is slack even when the sender chooses her most preferred signal $\pi_{01} = 0$, $\pi_{10} = 1$, in which case she persuades the receiver with certainty.

3 Optimal signal when the receiver is naive

In this section and Sect. 4 we assume that the receiver is naive. More precisely, while the receiver knows $(\mu(\sigma_0), \mu(\sigma_1))$ and $(\mu(\rho_0), \mu(\rho_1))$, he perceives σ and ρ as independent random variables. This gives rise to the misspecified prior $\hat{\mu}$ about Ω, where the receiver assumes that each element of the support of $\hat{\mu}$ is the product of corresponding marginals, i.e. $\hat{\mu}_{k\ell} = \mu^{\sigma}_{k}\mu^{\rho}_{\ell}$. We present the misspecified prior in the following Table 4 with marginal distributions written on the sides.

We call such a misspecified prior $\hat{\mu}$ a *simplistic worldview*. We assume that the sender's prior μ and receiver's prior $\hat{\mu}$ are common knowledge, i.e. they disagree on the prior. Moreover, while both the sender and receiver believe that they hold the correct prior, we suppose that the sender's prior is objectively correct. Lemma 3 below demonstrates that the simplistic worldview relates to the correct prior in a systematic way.

Lemma 3 Let $\mu \in \Delta(\Omega)$ and $k, \ell \in \{0, 1\}$. Let

$$\hat{\mu}_{kk} - \mu_{kk} \equiv c.$$

In this case,

$$\hat{\mu}_{k\ell} - \mu_{k\ell} = -c, \text{ for } k \neq \ell.$$

Lemma 3 shows that the amount of change in the priors about the states (due to the simplistic worldview) in which preferences are aligned and misaligned is the same but in opposite directions. That is, if the naive receiver perceives the states in which preferences are aligned as more likely than the rational receiver (i.e. if $c > 0$), then he perceives the states in which preferences are misaligned as less likely than the rational receiver, by the same amount (and vice versa) (Table 5). Therefore, we can represent the misspecified prior $\hat{\mu}$ in terms of the correct prior μ.

The naive receiver updates his belief according to the signal chosen by the sender and the misspecified prior $\hat{\mu}$. Let $\hat{\mu}^m(\rho_\ell)$ denote the posterior belief about state ρ_ℓ when the receiver updates according to $\hat{\mu}$. Thus, the receiver's posterior that the state is ρ_1 upon observing message 1 is

$$\hat{\mu}^1(\rho_1) = \frac{\hat{\mu}_{01}\pi_{01} + \hat{\mu}_{11}\pi_{11}}{\hat{\mu}_{00}\pi_{00} + \hat{\mu}_{10}\pi_{10} + \hat{\mu}_{01}\pi_{01} + \hat{\mu}_{11}\pi_{11}}. \tag{3}$$

Table 4 Misspecified prior $\hat{\mu}$

(Σ, R)	ρ_0	ρ_1	
σ_0	$\mu(\sigma_0)\mu(\rho_0)$	$\mu(\sigma_0)\mu(\rho_1)$	$\mu(\sigma_0)$
σ_1	$\mu(\sigma_1)\mu(\rho_0)$	$\mu(\sigma_1)\mu(\rho_1)$	
	$\mu(\rho_0)$		

Table 5 $\hat{\mu}$ expressed in terms of μ.

(Σ, R)	ρ_0	ρ_1	
σ_0	$\mu_{00} + c$	$\mu_{01} - c$	$\mu(\sigma_0)$
σ_1	$\mu_{10} - c$	$\mu_{11} + c$	
	$\mu(\rho_0)$		

While the sender calculates her expected utility according to μ, she knows that the naive receiver updates according to $\hat{\mu}$. Therefore, the straightforwardness constraint in the sender's optimization problem is now given by $\hat{\mu}_{00}\pi_{00} + \hat{\mu}_{10}\pi_{10} \leq \hat{\mu}_{01}\pi_{01} + \hat{\mu}_{11}\pi_{11}$. Let $\hat{\pi}^*$ denote the optimal signal when the receiver is naive. Note that in $\hat{\pi}^*$, similar to the rational receiver case, the sender's optimal choice in states (σ_0, ρ_0) and (σ_1, ρ_1) relaxes the straightforwardness constraint. Thus, $\hat{\pi}^*_{00} = 0$ and $\hat{\pi}^*_{11} = 1$. This allows us to simplify the sender's problem:

$$\begin{aligned} \max_{\pi \in \Pi} \ & \mu_{10}\pi_{10} - \mu_{01}\pi_{01} \\ & s.t. \quad \hat{\mu}_{10}\pi_{10} - \hat{\mu}_{01}\pi_{01} \leq \hat{\mu}_{11}. \end{aligned} \tag{4}$$

Similar to the result we demonstrated in Lemma 2, when action 1 is recommended under $\hat{\pi}^*$, this is also the sender's preferred action. The formal result and its proof can be found in Lemma A2 in the appendix.

The next proposition provides the full characterization of an optimal signal for a naive receiver.

Proposition 2 Let the priors of the sender and receiver be μ and $\hat{\mu}$, respectively. The optimal signal is given by

$$(\hat{\pi}^*_{00}, \hat{\pi}^*_{01}, \hat{\pi}^*_{10}, \hat{\pi}^*_{11}) = \begin{cases} (0, 0, 1, 1) \ \text{ if } \ \hat{\mu}_{11} \geq \hat{\mu}_{10}, \\ (0, \beta, \gamma, 1) \ \text{ otherwise,} \end{cases}$$

where

$$(\beta, \gamma) = \begin{cases} \left(0, \frac{\hat{\mu}_{11}}{\hat{\mu}_{10}}\right) & \text{if } \frac{\hat{\mu}_{01}}{\hat{\mu}_{10}} < \frac{\mu_{01}}{\mu_{10}}, \\ \left(1, \frac{\hat{\mu}_{01}+\hat{\mu}_{11}}{\hat{\mu}_{10}}\right) & \text{if } \frac{\hat{\mu}_{01}}{\hat{\mu}_{10}} > \frac{\mu_{01}}{\mu_{10}} \text{ and } \frac{\hat{\mu}_{01}+\hat{\mu}_{11}}{\hat{\mu}_{10}} \leq 1, \\ \left(\frac{\hat{\mu}_{10}-\hat{\mu}_{11}}{\hat{\mu}_{01}}, 1\right) & \text{if } \frac{\hat{\mu}_{01}}{\hat{\mu}_{10}} > \frac{\mu_{01}}{\mu_{10}} \text{ and } \frac{\hat{\mu}_{01}+\hat{\mu}_{11}}{\hat{\mu}_{10}} \geq 1, \end{cases}$$

and $\hat{\mu}_{10}\gamma - \hat{\mu}_{01}\beta = \hat{\mu}_{11}$ if $\hat{\mu}_{01}/\hat{\mu}_{10} = \mu_{01}/\mu_{10}$.

The optimal signal in Proposition 2 has several features similar to the one in Proposition 1.[11] First, the sender always recommends switching the default action in state (σ_1, ρ_1) (i.e. $\hat{\pi}^*_{11} = 1$) and never recommends switching in state (σ_0, ρ_0) (i.e. $\hat{\pi}^*_{00} = 0$). Second, when $\hat{\mu}_{11} \geq \hat{\mu}_{10}$, the sender can ensure that the receiver takes her desired action with probability 1 (i.e. $\hat{\pi}^*_{01} = 0$ and $\hat{\pi}^*_{10} = 1$). On the other hand, if $\hat{\mu}_{11} < \hat{\mu}_{10}$ then $\hat{\pi}^*_{01}$ and $\hat{\pi}^*_{10}$ are pinned down by the relation between the slope of the straightforwardness constraint $(\hat{\mu}_{01}/\hat{\mu}_{10})$ and the slope of the objective function (μ_{01}/μ_{10}).

One very important difference from Proposition 1 is that when $\mu_{10} \neq \mu_{01}$ (which implies $\hat{\mu}_{01}/\hat{\mu}_{10} \neq \mu_{01}/\mu_{10}$) the optimal signal in Proposition 2 is *unique*. This is a direct consequence of the receiver's naivete, as it leads to the sender and receiver having different priors. More precisely, if they have the same prior then the sender is indifferent between marginally increasing π_{10} or marginally decreasing π_{01}. On the other hand, if they have different priors then this indifference no longer holds as the two options tighten the straightforwardness constraint to a different extent. Hence, the sender chooses the option that tightens the straightforwardness constraint less as to optimally persuade the receiver.

4 Welfare implications of a simplistic worldview

In this section, we analyze a number of general properties of the sender's problem when the receiver is naive in comparison to the benchmark (2). This allows us to characterize the welfare implications of the receiver's naivete and explain the economic intuition behind them. Lemma 4 demonstrates how the receiver's naivete affects the sender's ability to manipulate the receiver's actions.

Lemma 4 Assume that $c > 0$. For any $\pi \in \Pi$ with $\pi_{00} = 0$ and $\pi_{11} = 1$ it holds that

(i) $\hat{\mu}^1(\rho_1) > \mu^1(\rho_1)$. That is, recommendation 1 induces a higher belief about state ρ_1 for a receiver with a simplistic worldview than for a rational receiver.
(ii) $(\hat{\mu}_{11}/\hat{\mu}_{10}) + (\hat{\mu}_{01}/\hat{\mu}_{10})\pi_{01} > (\mu_{11}/\mu_{10}) + (\mu_{01}/\mu_{10})\pi_{01}$. That is, the straightforwardness constraint in (4) is more relaxed relative to the benchmark (2).

If $c < 0$, the converse of the inequalities in (i) and (ii) hold.

The intuition of Lemma 4 is as follows. First, if $c > 0$ then by Lemma 3 the naive receiver perceives the aligned states (σ_0, ρ_0) and (σ_1, ρ_1) as more likely (and the two misaligned states as less likely) than the rational receiver. Hence, the naive receiver associates recommendation 1 with ρ_1 more than the rational receiver and thus has a higher posterior belief.

[11] It is important to note that our characterization of the optimal signal holds for an *arbitrary* prior of the receiver. Nevertheless, we are particularly interested in the simplistic worldview of the receiver and in its welfare implications. We consider comparative statics with respect to an arbitrary prior in Sect. 5.

Part (*ii*) of the lemma follows easily from (*i*). Since the sender can induce a posterior belief of at least 1/2 that the state is ρ_1 when the receiver is naive by recommending 1 more frequently than before, the straightforwardness constraint in (4) is more relaxed relative to the benchmark (2). It is easy to see that the converse of the above argument hold if $c < 0$.

Importantly, Lemma 4 has direct implications on the sender's and receiver's welfare as it pins down the conditions under which it is "easier" for the sender to persuade a naive receiver. In particular, if $c > 0$ then by (*ii*) the straightforwardness constraint is more slack than in the benchmark. Hence, the sender can decrease π_{01} and increase π_{10} so that the receiver chooses the sender-preferred actions in the misaligned states (σ_0, ρ_1) and (σ_1, ρ_0) with a higher probability. Clearly, this benefits the sender. Moreover, since total welfare is the same at the optimum in both cases, it follows that naivete harms the receiver. Intuitively, converse arguments hold if $c < 0$.

The welfare implications of the receiver's simplistic worldview are summarized in Proposition 3.

Proposition 3 Let $c > 0$. Then,

(i) the receiver's simplistic worldview harms him and benefits the sender and
(ii) the differences in their expected utilities are strict if and only if $\mu_{11} < \mu_{10}$.

If $c < 0$, then the converse of (*i*) holds and the differences in their expected utilities are strict if and only if $\hat{\mu}_{11} < \hat{\mu}_{10}$.

Part (*ii*) of Proposition 3 follows from the structure of the optimal signals. If the sender is facing a rational receiver, then she can ensure that the receiver takes her preferred action whenever $\mu_{11} \geq \mu_{10}$. Hence, in this case if $c > 0$ the sender can only weakly benefit from the receiver being naive. Otherwise, the sender can always guarantee a strict improvement. On the other hand, if $c < 0$ and the sender is facing a naive receiver, then she can ensure that the receiver takes her preferred action whenever $\hat{\mu}_{11} \geq \hat{\mu}_{10}$. But since the receiver's naivete is detrimental for the sender if $c < 0$, it follows that the sender obtains her preferred outcome with certainty also when the receiver is rational. In this case, the sender is only weakly harmed by the receiver's naivete.

Proposition 3 can also be interpreted in terms of the receiver's misperception of ex ante preference alignment. Recall that by Lemma 3 if $c > 0$ then the naive receiver perceives the misaligned states as less likely than the rational receiver. Therefore, Lemma 4 implies that in this case it is easier for the sender to persuade the receiver. Hence, *naivete makes the receiver more manipulable and is detrimental for the receiver*. Intuitively, if $c < 0$ the naivete of the receiver constrains the sender more and thus the "correct" recommendation is sent with a higher probability. In this case, *naivete makes the receiver less manipulable and is beneficial for the receiver*.

5 Comparative statics on the receiver's prior

The optimal probabilities of recommending action 1 in the two states of preference alignment does not depend on the receiver's prior: it follows from Proposition 1 and Proposition 2 that the sender always recommends action 1 in state (σ_1, ρ_1) and action 0 in state (σ_0, ρ_0). Yet, in the two states of preference misalignment, (σ_0, ρ_1) and (σ_1, ρ_0), the sender's ability to beneficially fine-tune π_{01} and π_{10} crucially depends on the receiver's prior. In this section, rather than focusing on the receiver's simplistic worldview, we consider the sender's problem when the receiver has an arbitrary prior $\tilde{\mu} \in \Delta(\Omega)$, and denote the sender's optimal choice by $\tilde{\pi}^*$. We show how the receiver's prior $\tilde{\mu}$ affects the choice of $\tilde{\pi}^*_{01}$ and $\tilde{\pi}^*_{10}$, and thus the sender's ability to persuade the receiver.

Observe that lower the receiver's prior about state (σ_1, ρ_0), higher the probability the sender can assign to π_{10} in an optimal signal. It follows that if the receiver attaches 0 probability to state (σ_1, ρ_0), i.e. $\tilde{\mu}_{10} = 0$, then the sender can recommend action 1 with probability 1. Similar arguments hold for state (σ_0, ρ_1). Therefore, $\tilde{\mu}_{10} = \tilde{\mu}_{01} = 0$ is a sufficient condition on the receiver's prior for the sender to induce her preferred actions with probability 1 by choosing $(\tilde{\pi}^*_{00}, \tilde{\pi}^*_{01}, \tilde{\pi}^*_{10}, \tilde{\pi}^*_{11}) = (0, 0, 1, 1)$.[12]

Intuitively, when the receiver attaches 0 probability to the states in which preferences are misaligned, the receiver perceives preferences as perfectly aligned, and is thus always willing to follow the sender's recommendations.

Similarly and intuitively, a lower prior belief about state (σ_1, ρ_1) makes the sender's job more difficult; it is harder to persuade a receiver who thinks the state in which the sender and receiver both prefer action 1 is less likely. Hence, this captures a change in the receiver's prior that unambiguously restricts the sender's ability to persuade the receiver. The next result formalizes this observation.

Proposition 4 Let $\tilde{\mu} \in \Delta(\Omega)$. Then either (i) $\tilde{\pi}^*_{01}$ is weakly decreasing in $\tilde{\mu}_{11}$ or (ii) $\tilde{\pi}^*_{10}$ is weakly increasing in $\tilde{\mu}_{11}$. If $\tilde{\mu}_{11} < \tilde{\mu}_{10}$, then the relations in (i) and (ii) are strict.

The intuition for Proposition 4 is simple: as the likelihood of state (σ_1, ρ_1) decreases, the sender has to adjust the optimal signal such that the recommendations that match the receiver's preference (i.e. recommendation 1 in states (σ_1, ρ_1) and (σ_0, ρ_1) and recommendation 0 in the two remaining states) are sent with a higher probability. This can happen in two ways: send recommendation 1 (i) less often when the state is (σ_1, ρ_0) or (ii) more frequently when the state is (σ_0, ρ_1). That is, to respect the straightforwardness constraint, the sender has to mix the recommendations in a manner sufficiently more favorable for the receiver.

[12] This implies that if the sender had the option to alter the prior of the receiver without a cost before the communication phase, she would decrease the priors of the states in which preferences are misaligned as much as possible.

6 Conclusion

Real-world examples of Bayesian persuasion problems, such as corporate consulting, often involve the sender and receiver caring about different, but connected issues of a problem. This paper studies the impact of the receiver's naivete regarding the interconnected nature of the two dimensions of the state of the world on the optimal outcomes for both parties. First, as a benchmark, we study the case of a rational receiver, i.e., the sender and receiver both know the correct prior distribution of the states. We show that regardless of the receiver being rational or naive, it is optimal for the sender to truthfully disclose the states whenever her preference is aligned with the receiver. On the other hand, optimal disclosure in states where preferences are misaligned crucially depends on the receiver's perception of the prior.

Our findings reveal that the receiver's naivete can disadvantage him while favoring the sender, particularly when it leads the receiver to overestimate the likelihood of states where both parties prefer the same action and underestimate states where they have conflicting preferences. Intuitively, this makes the receiver "easier" to persuade; the sender can commit to sending the "wrong" message with a higher probability. Conversely, if the naivete leads the receiver to perceive aligned states as less likely and misaligned states as more probable, the receiver benefits at the expense of the sender.

In this study, for clarity of exposition we use a framework with two actions and four states rather than general finite action and state spaces. Extending our model to a general finite state space in both the sender's and receiver's dimensions can be handled using our approach, however, it would complicate the analysis without yielding substantial new insights to the qualitative features of our results. On the other hand, allowing for a general finite action space goes beyond the analytic approach that we take and requires considering a different model, which constitutes an avenue for further research.

Proofs

Proof of Lemma 1 For the straightforwardness constraint to be satisfied, it must hold that $\mu^1(\rho_1) \geq 1/2$, which is equivalent to

$$\frac{\mu_{01}\pi_{01} + \mu_{11}\pi_{11}}{\mu_{01}\pi_{01} + \mu_{11}\pi_{11} + \mu_{10}\pi_{10} + \mu_{00}\pi_{00}} \geq \frac{1}{2}.$$

Rearranging the inequality yields the straightforwardness constraint. □

Lemma A1 Let $\pi \in \Pi$. If $\mu_{00}\pi_{00} + \mu_{10}\pi_{10} \leq \mu_{01}\pi_{01} + \mu_{11}\pi_{11}$, then $\mu_{01}(1 - \pi_{01}) + \mu_{11}(1 - \pi_{11}) < \mu_{00}(1 - \pi_{00}) + \mu_{10}(1 - \pi_{10})$.

Proof Let $\mu_{00}\pi_{00} + \mu_{10}\pi_{10} \leq \mu_{01}\pi_{01} + \mu_{11}\pi_{11}$, which implies $-\mu_{01}\pi_{01} - \mu_{11}\pi_{11} \leq -\mu_{00}\pi_{00} - \mu_{10}\pi_{10}$. Recall that $\mu(\rho_1) < \mu(\rho_0)$, i.e. $\mu_{01} + \mu_{11} < \mu_{00} + \mu_{10}$. Hence, it follows that

$$\mu_{01} + \mu_{11} - \mu_{01}\pi_{01} - \mu_{11}\pi_{11} < \mu_{00} + \mu_{10} - \mu_{00}\pi_{00} - \mu_{10}\pi_{10}.$$

Rewriting the above inequality provides the desired result. Hence, the straightforwardness constraint for 0 is satisfied. That is,

$$\mu^0(\rho_0) > \frac{\mu_{10}(1-\pi_{10}) + \mu_{00}(1-\pi_{00})}{\mu_{10}(1-\pi_{10}) + \mu_{00}(1-\pi_{00}) + \mu_{01}(1-\pi_{01}) + \mu_{11}(1-\pi_{11})} > \frac{1}{2}.$$

□

Proof of Proposition 1 We solve the following problem:

$$\begin{aligned} \max_{\pi\in\Pi}\ & \mu_{11}\pi_{11} + \mu_{10}\pi_{10} - \mu_{01}\pi_{01} - \mu_{00}\pi_{00} \\ & \text{s.t. } \mu_{00}\pi_{00} + \mu_{10}\pi_{10} \leq \mu_{01}\pi_{01} + \mu_{11}\pi_{11}. \end{aligned} \tag{5}$$

First suppose that $\mu_{10} \leq \mu_{11}$. It is clear that setting $\pi^* = (\pi^*_{00}, \pi^*_{01}, \pi^*_{10}, \pi^*_{11}) = (0, 0, 1, 1)$ satisfies the straightforwardness constraint in (5) and maximizes the sender's expected utility since it provides the upper bound payoff of 1. Hence, π^* is the unique solution to the sender's optimization problem.

Now suppose that $\mu_{10} > \mu_{11}$. First observe that it is optimal in this case to set $\pi^*_{11} = 1$ and $\pi^*_{00} = 0$ as well, since it increases the sender's expected utility and relaxes the straightforwardness constraint. Hence, substituting $\pi^*_{11} = 1$ and $\pi^*_{00} = 0$ into the sender's optimization problem and simplifying we obtain

$$\begin{aligned} \max_{\pi\in\Pi}\ & \mu_{10}\pi_{10} - \mu_{01}\pi_{01} \\ & \text{s.t. } \mu_{10}\pi_{10} \leq \mu_{01}\pi_{01} + \mu_{11}. \end{aligned} \tag{6}$$

It is easy to see from (6) that the objective function is increasing in π_{10} and decreasing in π_{01}. Hence, if (6) is not binding then the sender can marginally increase π_{10} or decrease π_{01} to achieve a higher expected payoff. Thus, the constraint is binding at the optimum.

Therefore, since both the objective function and the constraint are linear (in π_{10} and π_{01}), there exists a continuum of solutions to (6). □

Proof of Lemma 2 The sender prefers action $a = 1$ if $\mu^1(\sigma_1) \geq \frac{1}{2}$, which is equivalent to

$$\begin{aligned} \frac{\mu_{10}\pi_{10} + \mu_{11}\pi_{11}}{\mu_{10}\pi_{10} + \mu_{11}\pi_{11} + \mu_{01}\pi_{01} + \mu_{00}\pi_{00}} &\geq \frac{1}{2} \\ \iff \mu_{10}\pi_{10} + \mu_{11}\pi_{11} &\geq \mu_{01}\pi_{01} + \mu_{00}\pi_{00}. \end{aligned} \tag{7}$$

To proceed, we demonstrate that the signal solving (5) satisfies the alignment constraint (7). First, consider the case $\mu_{11} \geq \mu_{10}$, then imputing π^* in alignment constraint yields

$$0 \leq \mu_{10} + \mu_{11},$$

so it is satisfied. Second, consider the case $\mu_{11} < \mu_{10}$, then the straightforwardness constraint is binding,

$$\mu_{10}\pi^*_{10} - \mu_{01}\pi^*_{01} = \mu_{11},$$

while the alignment constraint (7) can be written as

$$-\left(\mu_{10}\pi^*_{10} - \mu_{01}\pi^*_{01}\right) \leq \mu_{11},$$

and this inequality is implied by the binding straightforwardness constraint. Thus, π^* satisfies (7). □

Proof of Lemma 3 Let $\hat{\mu}_{kk} - \mu_{kk} = c$ and assume that $k \neq \ell$. Then,

$$\begin{aligned}\hat{\mu}_{k\ell} - \mu_{k\ell} = \mu(\sigma_k)(1 - \mu(\rho_k)) - \mu_{k\ell} &= \mu_{k\ell} + \mu_{kk} - \mu(\sigma_k)\mu(\rho_k) - \mu_{k\ell} \\ &= \mu(\sigma_k)\mu(\rho_k) - c - \mu(\sigma_k)\mu(\rho_k) = -c.\end{aligned}$$

□

Proof of Proposition 2 We solve the problem

$$\begin{aligned}\max_{\pi \in \Pi}\ & \mu_{10}\pi_{10} - \mu_{01}\pi_{01} \\ s.t.\quad & \hat{\mu}_{10}\pi_{10} - \hat{\mu}_{01}\pi_{01} \leq \hat{\mu}_{11}.\end{aligned}$$

As the problem is a linear program, the solution lies either on the straightforwardness constraint or on one of the boundaries of the set Π. First, suppose that $\hat{\mu}_{11} \geq \hat{\mu}_{10}$. Then the solution lies on the boundary and the straightforwardness constraint is inactive for $\hat{\mu}_{11} > \hat{\mu}_{10}$. It follows easily that $\pi^*_{01} = 0, \pi^*_{10} = 1$.

Next, suppose that $\hat{\mu}_{11} < \hat{\mu}_{10}$. All three possible types of solutions for this case are presented in Fig. 1. We proceed with characterizing each of the three cases.

Case (a) The active constraints are $\pi_{01} \geq 0$ and $\hat{\mu}_{10}\pi_{10} - \hat{\mu}_{01}\pi_{01} \leq \hat{\mu}_{11}$. Hence, $\hat{\pi}^*_{01} = 0$ and $\hat{\pi}^*_{10}$ solves $\hat{\mu}_{10}\pi_{10} - \hat{\mu}_{01}\pi_{01} = \hat{\mu}_{11}$, i.e. $\pi^*_{10} = \hat{\mu}_{11}/\hat{\mu}_{10}$. The solution to (4) has this particular form if and only if the slope of the objective function is steeper than the slope of the straightforwardness constraint, i.e. $(\hat{\mu}_{01}/\hat{\mu}_{10}) < (\mu_{01}/\mu_{10})$.

Case (b) The active constraints are $\pi_{01} \leq 1$ and $\hat{\mu}_{10}\pi_{10} - \hat{\mu}_{01}\pi_{01} \leq \hat{\mu}_{11}$. Hence, $\hat{\pi}^*_{01} = 1$ and $\hat{\pi}^*_{10}$ solves $\hat{\mu}_{10}\pi_{10} - \hat{\mu}_{01}\pi_{01} = \hat{\mu}_{11}$, i.e. $\hat{\pi}^*_{10} = (\hat{\mu}_{01} + \hat{\mu}_{11})/\hat{\mu}_{10}$. The solution to (4) has this form if and only if (i) $(\hat{\mu}_{01}/(\hat{\mu}_{10}) > (\mu_{01}/\mu_{10})$ (i.e. the slope of the objective function is flatter than the slope of the straightforwardness constraint) and (ii) $(\hat{\mu}_{01} + \hat{\mu}_{11})/\hat{\mu}_{10} \leq 1$ (which implies that the constraint $\pi_{10} \leq 1$ is inactive).

Case (c) The active constraints are $\pi_{10} \leq 1$ and $\hat{\mu}_{10}\pi_{10} - \hat{\mu}_{01}\pi_{01} \leq \hat{\mu}_{11}$. Hence, $\hat{\pi}^*_{10} = 1$ and $\hat{\pi}^*_{01}$ solves $\hat{\mu}_{10}\pi_{10} - \hat{\mu}_{01}\pi_{01} = \hat{\mu}_{11}$, i.e. $\hat{\pi}^*_{01} = (\hat{\mu}_{10} - \hat{\mu}_{11})/\hat{\mu}_{01}$.

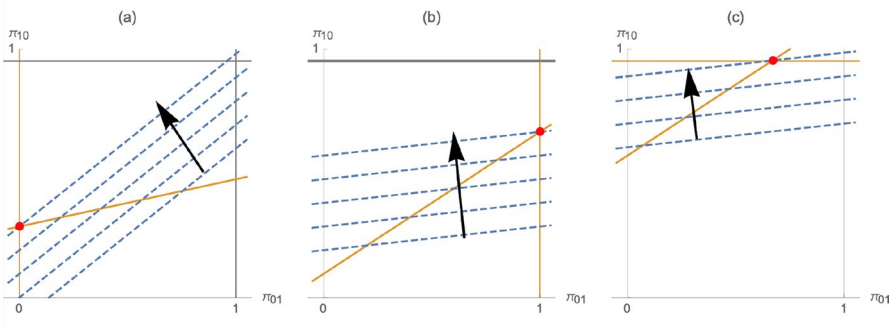

Fig. 1 Three possible forms of solutions to the problem (4). Indifference curves are in blue and dashed. Active constraints are in orange

The solution to (4) has this form if and only if (i) $\hat{\mu}_{01}/\hat{\mu}_{10} > \mu_{01}/\mu_{10}$, and (ii) $(\hat{\mu}_{01} + \hat{\mu}_{11})/\hat{\mu}_{10} \geq 1$ (which implies that the constraint $\pi_{10} \leq 1$ is active).

Finally, if $(\hat{\mu}_{01}/\hat{\mu}_{10}) = (\mu_{01}/\mu_{10})$, then any π_{10} and π_{01} such that $\hat{\mu}_{10}\pi_{10} - \hat{\mu}_{01}\pi_{01} = \hat{\mu}_{11}$ solves (4). □

Lemma A2 Given $\hat{\pi}^* \in \Pi$ that solves (4), it holds that $\mu^1(\sigma_1) \geq 1/2$.

Proof Here, we demonstrate that the optimal signal $\hat{\pi}^*$ satisfies the alignment constraint for $a = 1$ presented in (7). First, assume $\hat{\mu}_{11} \geq \hat{\mu}_{10}$, then imputing π^* in alignment constraint yields

$$0 \leq \mu_{10} + \mu_{11},$$

which holds for any μ_{10} and μ_{11}.

Now assume $\hat{\mu}_{11} < \hat{\mu}_{10}$. Thus, we have cases (a), (b), and (c) by Proposition 2. Consider case (a), i.e. $\frac{\hat{\mu}_{01}}{\hat{\mu}_{10}} < \frac{\mu_{01}}{\mu_{10}}$. Imputing $\hat{\pi}^*_{01} = 0$ in the alignment constraint yields

$$0 \leq \mu_{10}\pi_{10} + \mu_{11} \iff \pi_{10} \geq -\frac{\mu_{11}}{\mu_{10}},$$

which trivially holds.

Consider case (b), i.e. $\frac{\hat{\mu}_{01}}{\hat{\mu}_{10}} > \frac{\mu_{01}}{\mu_{10}}$ and $\frac{\hat{\mu}_{01}+\hat{\mu}_{11}}{\hat{\mu}_{10}} \leq 1$. Imputing $\hat{\pi}^*_{01} = 1$ in the alignment constraint yields

$$\mu_{01} \leq \mu_{10}\pi_{10} + \mu_{11} \iff \pi_{10} \geq \frac{\mu_{01} - \mu_{11}}{\mu_{10}},$$

while $\hat{\pi}^*_{10} = \frac{\hat{\mu}_{01}+\hat{\mu}_{11}}{\hat{\mu}_{10}}$. Further, it holds that $\frac{\hat{\mu}_{01}+\hat{\mu}_{11}}{\hat{\mu}_{10}} \geq \frac{\mu_{01}-\mu_{11}}{\mu_{10}}$. To see that note

$$\frac{\hat{\mu}_{01} + \hat{\mu}_{11}}{\hat{\mu}_{10}} \geq \frac{\mu_{01} - \mu_{11}}{\mu_{10}} \iff \frac{\hat{\mu}_{01}}{\hat{\mu}_{10}} - \frac{\mu_{01}}{\mu_{10}} \geq -\left(\frac{\hat{\mu}_{11}}{\hat{\mu}_{10}} + \frac{\mu_{11}}{\mu_{10}}\right),$$

which holds given the case (b) parametric restrictions.

Consider case (c), i.e. $\frac{\hat{\mu}_{01}}{\hat{\mu}_{10}} > \frac{\mu_{01}}{\mu_{10}}$ and $\frac{\hat{\mu}_{01}+\hat{\mu}_{11}}{\hat{\mu}_{10}} \geq 1$. Imputing $\hat{\pi}^*_{10} = 1$ in the alignment constraint yields

$$\mu_{01}\pi_{01} \leq \mu_{10} + \mu_{11} \iff \pi_{01} \leq \frac{\mu_{10} + \mu_{11}}{\mu_{01}},$$

while $\hat{\pi}^*_{01} = \frac{\hat{\mu}_{10}-\hat{\mu}_{11}}{\hat{\mu}_{01}}$. Further, it holds that $\frac{\hat{\mu}_{10}-\hat{\mu}_{11}}{\hat{\mu}_{01}} \leq \frac{\mu_{10}+\mu_{11}}{\mu_{01}}$. To see that note

$$\frac{\hat{\mu}_{10} - \hat{\mu}_{11}}{\hat{\mu}_{01}} \leq \frac{\mu_{10} + \mu_{11}}{\mu_{01}} \iff \frac{\hat{\mu}_{10}}{\hat{\mu}_{01}} - \frac{\mu_{10}}{\mu_{01}} \leq \frac{\hat{\mu}_{11}}{\hat{\mu}_{01}} + \frac{\mu_{11}}{\mu_{01}},$$

which trivially holds as the case (c) parametric restrictions imply that the left-hand side of the last inequality is negative.

Finally, consider the non-generic case $\hat{\mu}_{01}/\hat{\mu}_{10} = \mu_{01}/\mu_{10}$. In this case, the straightforwardness constraint is binding at optimum, $\hat{\mu}_{10}\hat{\pi}^*_{10} - \hat{\mu}_{01}\hat{\pi}^*_{01} = \hat{\mu}_{11}$. From the binding straightforwardness constraint and the non-genericity condition, it follows that

$$\frac{\hat{\mu}_{10}}{\hat{\mu}_{01}} = \frac{\hat{\mu}_{11}}{\hat{\mu}_{01}} \cdot \frac{1}{\pi_{10}} + \frac{\pi_{01}}{\pi_{10}} = \frac{\mu_{10}}{\mu_{01}}. \tag{8}$$

Rewriting the alignment constraint

$$\begin{aligned} \mu_{01}\pi_{01} - \mu_{10}\pi_{10} \leq \mu_{11} &\iff \mu_{10}\pi_{10} + \mu_{11} \geq \mu_{01}\pi_{01} \\ &\iff \frac{\mu_{10}}{\mu_{01}} \geq \frac{\pi_{01}}{\pi_{10}} - \frac{\mu_{11}}{\mu_{01}} \cdot \frac{1}{\pi_{10}}, \end{aligned}$$

and plugging (8) gives

$$\frac{\hat{\mu}_{11}}{\hat{\mu}_{01}} \cdot \frac{1}{\pi_{10}} + \frac{\pi_{01}}{\pi_{10}} \geq \frac{\pi_{01}}{\pi_{10}} - \frac{\mu_{11}}{\mu_{01}} \cdot \frac{1}{\pi_{10}} \iff \frac{\hat{\mu}_{11}}{\hat{\mu}_{01}} \geq -\frac{\mu_{11}}{\mu_{01}},$$

which trivially holds.

Hence, given $\hat{\pi}^*$, (7) is satisfied. □

Proof of Lemma 4 We begin by proving part (*i*). Since it is optimal to set $\hat{\pi}^*_{00} = 0$ and $\hat{\pi}^*_{11} = 1$, (3) can be rewritten as

$$\hat{\mu}^1(\rho_1) = \frac{\hat{\mu}_{01}\pi_{01} + \hat{\mu}_{11}}{\hat{\mu}_{01}\pi_{01} + \hat{\mu}_{11} + \hat{\mu}_{10}\pi_{10}}.$$

By Lemma 3, it follows that

$$\hat{\mu}^1(\rho_1) = \frac{(\mu_{01} - c)\pi_{01} + (\mu_{11} + c)}{(\mu_{01} - c)\pi_{01} + (\mu_{11} + c) + (\mu_{10} - c)\pi_{10}} = \frac{\mu_{01}\pi_{01} + \mu_{11} + c(1 - \pi_{01})}{\mu_{01}\pi_{01} + \mu_{11} + \mu_{10}\pi_{10} + c(1 - \pi_{01} - \pi_{10})}.$$

Suppose that $c > 0$. Then it holds that $c(1 - \pi_{01}) > c(1 - \pi_{01} - \pi_{10})$, and thus

$$\mu^1(\rho_1) = \frac{\mu_{01}\pi_{01} + \mu_{11}}{\mu_{01}\pi_{01} + \mu_{11} + \mu_{10}\pi_{10}} < \frac{\mu_{01}\pi_{01} + \mu_{11} + c(1 - \pi_{01})}{\mu_{01}\pi_{01} + \mu_{11} + \mu_{10}\pi_{10} + c(1 - \pi_{01} - \pi_{10})} = \hat{\mu}^1(\rho_1).$$

Next suppose that $c < 0$. Then it holds that $c(1 - \pi_{01}) < c(1 - \pi_{01} - \pi_{10})$, and thus

$$\mu^1(\rho_1) = \frac{\mu_{01}\pi_{01} + \mu_{11}}{\mu_{01}\pi_{01} + \mu_{11} + \mu_{10}\pi_{10}} > \frac{\mu_{01}\pi_{01} + \mu_{11} + c(1 - \pi_{01})}{\mu_{01}\pi_{01} + \mu_{11} + \mu_{10}\pi_{10} + c(1 - \pi_{01} - \pi_{10})} = \hat{\mu}^1(\rho_1).$$

Now we prove part (*ii*). Recall the sender's simplified problems (2) for the rational receiver and (4) for the naive receiver.

Rewriting the constraints in (2) and (4), respectively, yields

$$\pi_{10} \leq \frac{\mu_{01}}{\mu_{10}}\pi_{01} + \frac{\mu_{11}}{\mu_{10}}, \quad \pi_{10} \leq \frac{\hat{\mu}_{01}}{\hat{\mu}_{10}}\pi_{01} + \frac{\hat{\mu}_{11}}{\hat{\mu}_{10}}.$$

Define $g_1(\pi_{01}) \equiv (\mu_{01}/\mu_{10})\pi_{01} + (\mu_{11}/\mu_{10})$ and $g_2(\pi_{01}) \equiv (\hat{\mu}_{01}/\hat{\mu}_{10})\pi_{01} + (\hat{\mu}_{11}/\hat{\mu}_{10})$. By Lemma 3 it follows that

$$g_2(\pi_{01}) = \frac{\mu_{01} - c}{\mu_{10} - c}\pi_{01} + \frac{\mu_{11} + c}{\mu_{10} - c}.$$

Consider $g_1(\cdot)$ and $g_2(\cdot)$ at the boundaries of the set of feasible signals: $g_1(0) = (\mu_{11}/\mu_{10})$, $g_2(0) = (\mu_{11} + c)/(\mu_{10} - c)$ and $g_1(1) = (\mu_{01} + \mu_{11})/\mu_{10}$, $g_2(1) = (\mu_{01} + \mu_{11})/(\mu_{10} - c)$.

First suppose that $c > 0$. It follows that $g_1(0) < g_2(0)$ and $g_1(1) < g_2(1)$. Moreover, by linearity of $g_1(\cdot)$ and $g_2(\cdot)$, for all $\pi_{01} \in [0, 1]$ it holds that

$$g_1(\pi_{01}) < g_2(\pi_{01}).$$

Thus, the constraint in (2) is relatively more slack than in (4).

Next suppose that $c < 0$. It follows that $g_1(0) > g_2(0)$ and $g_1(1) > g_2(1)$. Again, by linearity of $g_1(\cdot)$ and $g_2(\cdot)$, for all $\pi_{01} \in [0, 1]$ it holds that

$$g_1(\pi_{01}) > g_2(\pi_{01}).$$

Thus, the constraint in (2) is relatively more tight than in (4). □

Proof of Proposition 3 Let V^* and $\hat{V}^*$ denote the sender's optimal (ex ante) expected payoff when the receiver is rational and when he is naive, respectively. Then,

$$V^* \equiv \mu_{11}\pi^*_{11} + \mu_{10}\pi^*_{10} + \mu_{01}(1-\pi^*_{01}) + \mu_{00}(1-\pi^*_{00}),$$
$$\hat{V}^* \equiv \mu_{11}\hat{\pi}^*_{11} + \mu_{10}\hat{\pi}^*_{10} + \mu_{01}(1-\hat{\pi}^*_{01}) + \mu_{00}(1-\hat{\pi}^*_{00}).$$

As the form of optimal signal in Propositions 1 and 2 depends on $sign(\mu_{11}-\mu_{10})$, we consider four cases based on $sign(\mu_{11}-\mu_{10})$ and $sign(c)$. First, assume $\mu_{11} \geq \mu_{10}$. If $c \geq 0$, then from Proposition 1 and 2 and Lemma 3, $(\pi^*_{00}, \pi^*_{01}, \pi^*_{10}, \pi^*_{11}) = (\hat{\pi}^*_{00}, \hat{\pi}^*_{01}, \hat{\pi}^*_{10}, \hat{\pi}^*_{11}) = (0,0,1,1)$. Thus, $V^* = \hat{V}^* = 1$. If $c < 0$, then either $\hat{\mu}_{11} \geq \hat{\mu}_{10}$ and, as above, $V^* = \hat{V}^* = 1$, or $\hat{\mu}_{11} < \hat{\mu}_{10}$ and (as, according to Proposition 2, the sender does not get her first best payoff) $\hat{V}^* < V^* = 1$. Thus, it follows that

$$\hat{V}^* \geq V^* \text{ if and only if } c \geq 0. \tag{9}$$

Second, assume $\mu_{11} < \mu_{10}$. By Proposition 1, the difference in the sender's optimal expected payoffs between the benchmark and the naive receiver case is given by

$$\begin{aligned} v \equiv \hat{V}^* - V^* &= \mu_{10}\left(\hat{\pi}^*_{10} - \pi^*_{10}\right) + \mu_{01}\left(\pi^*_{01} - \hat{\pi}^*_{01}\right) \\ &= \mu_{10}\left(\hat{\pi}^*_{10} - \gamma\right) + \mu_{01}\left(\frac{\mu_{10}}{\mu_{01}}\gamma - \frac{\mu_{11}}{\mu_{01}} - \hat{\pi}^*_{01}\right) \\ &= \mu_{10}\hat{\pi}^*_{10} - \mu_{01}\hat{\pi}^*_{01} - \mu_{11}. \end{aligned} \tag{10}$$

Plugging the optimal signal for case (a) from Proposition 2 in (10) yields

$$\begin{aligned} v &= \mu_{10}\frac{\hat{\mu}_{11}}{\hat{\mu}_{10}} - \mu_{11} = \mu_{10}\frac{\mu_{11}+c}{\hat{\mu}_{10}} - \mu_{11} \\ &= \frac{\mu_{10}\mu_{11} + \mu_{11}c - \mu_{10}\mu_{11} + \mu_{10}c}{\hat{\mu}_{10}} = \frac{c\left(\mu_{11}+\mu_{10}\right)}{\hat{\mu}_{10}}. \end{aligned} \tag{11}$$

Similarly, plugging the optimal signal for case (b) from Proposition 2 in (10) yields

$$\begin{aligned} v =& \mu_{10}\frac{\hat{\mu}_{01}+\hat{\mu}_{11}}{\hat{\mu}_{10}} - \mu_{01} - \mu_{11} \\ &= \mu_{10}\frac{\mu_{01} - c + \mu_{11} + c}{\hat{\mu}_{10}} - \mu_{01} - \mu_{11} = \frac{c\left(\mu_{11}+\mu_{01}\right)}{\hat{\mu}_{10}}. \end{aligned} \tag{12}$$

Finally, plugging the optimal signal for case (c) from Proposition 2 in (10) yields

$$\begin{aligned} \nu &= \mu_{10} - \mu_{01}\frac{\hat{\mu}_{10} - \hat{\mu}_{11}}{\hat{\mu}_{01}} - \mu_{11} \\ &= \frac{2c\mu_{01} + \big(\mu_{10} - \mu_{11}\big)\big(\hat{\mu}_{01}\big) - \mu_{01}\big(\mu_{10} - \mu_{11}\big)}{\hat{\mu}_{01}} = \frac{c\big(2\mu_{01} + \mu_{11} - \mu_{10}\big)}{\hat{\mu}_{01}} \\ &= \frac{c\big(\mu(\rho_1) - \mu_{10} + \mu_{01}\big)}{\hat{\mu}_{01}}. \end{aligned} \tag{13}$$

From the parametric condition for the case (c), $(\hat{\mu}_{01} + \hat{\mu}_{11})/\hat{\mu}_{10} \geq 1$ and thus it follows that $\mu(\rho_1) \geq \hat{\mu}_{10}$. By Lemma 3, this can be rewritten as $\mu(\rho_1) - \mu_{10} \geq -c$. Since we also have $\mu_{01} \geq c$ by Lemma 3, it holds that

$$\mu(\rho_1) - \mu_{10} + \mu_{01} \geq 0. \tag{14}$$

By (11), (12), (13), and (14) it follows that,

$$\hat{V}^* \geq V^* \text{ if and only if } c \geq 0. \tag{15}$$

Note that the sum of optimal expected payoffs of the sender and the receiver in the benchmark and the case of naive receiver is given by the same constant:

$$\begin{aligned} &\underbrace{\mu_{11}\pi^*_{11} + \mu_{00}(1 - \pi^*_{00}) + \mu_{10}\pi^*_{10} + \mu_{01}(1 - \pi^*_{01})}_{\text{the sender's expected payoff}} \\ &\quad + \underbrace{\mu_{11}\pi^*_{11} + \mu_{00}(1 - \pi^*_{00}) + \mu_{10}\big(1 - \pi^*_{10}\big) + \mu_{01}\pi^*_{01}}_{\text{the receiver's expected payoff}} \\ &= 2\big(\mu_{11}\pi^*_{11} + \mu_{00}(1 - \pi^*_{00})\big) + \mu_{10} + \mu_{01} \\ &= 2\big(\mu_{11} + \mu_{00}\big) + \mu_{10} + \mu_{01}, \end{aligned} \tag{16}$$

where the naive receiver's expected utility is evaluated using the correct (sender's) prior belief.

Denote the receiver's sender-optimal expected payoff in the benchmark and the case of naive receiver by U^* and $\hat{U}^*$, respectively. Since by (16) the sum of sender's and receiver's expected payoffs does not change between the benchmark and the case of naive receiver and the sender's expected payoff changes according to (9) and (15), it follows that

$$\hat{U}^* \leq U^* \text{ if and only if } c \geq 0, \tag{17}$$

i.e. when the receiver's naivete yields a higher expected payoff to the sender, it yields a lower expected payoff to the receiver.

Finally, given $c > 0$, by Lemma 4 the receiver's naivete relaxes the straightforwardness constraint. Thus, the difference between the expected payoffs in (9), (15), and (17) is strict if and only if in the benchmark case the sender does not obtain her first best payoff of 1, which is the case if and only if $\mu_{11} < \mu_{10}$. Similarly, given $c < 0$, by Lemma 4, the receiver's naivete tightens the straightforwardness constraint. Thus, the difference between the expected payoffs in (9), (15), and (17) is

strict if and only if the sender does not obtain her first-best payoff of 1 in the naive receiver case, which is true if and only if $\hat{\mu}_{11} < \hat{\mu}_{10}$. □

Proof of Proposition 4 Here, we use the fact that the result of Proposition 2 holds not only for the simplistic worldview prior $\hat{\mu}$, but for an arbitrary receiver's prior $\tilde{\mu}$. If $\tilde{\mu}_{11} \geq \tilde{\mu}_{10}$, then decreasing $\tilde{\mu}_{11}$ does not affect $\tilde{\pi}^*$. If $\tilde{\mu}_{11} < \tilde{\mu}_{10}$, then three cases are possible in line with Proposition 2.

In the first and second cases, it holds that $d\tilde{\pi}^*_{10}/d\tilde{\mu}_{11} > 0$. In the third case, it holds that $d\tilde{\pi}^*_{01}/d\tilde{\mu}_{11} < 0$. Finally, in the non-generic case of $\tilde{\mu}_{01}/\tilde{\mu}_{10} = \mu_{01}/\mu_{10}$, we have $\left(\tilde{\pi}^*_{01}, \tilde{\pi}^*_{10}\right) = \left(\tilde{\pi}^*_{01}, \frac{\tilde{\mu}_{11}}{\tilde{\mu}_{10}} + \frac{\tilde{\mu}_{01}}{\tilde{\mu}_{10}}\tilde{\pi}^*_{01}\right)$, and thus $d\tilde{\pi}^*_{10}/d\tilde{\mu}_{11} > 0$. □

Acknowledgements We would like to thank Ole Jann, Artyom Jelnov, Pavel Kocourek, Fedor Sandomirskiy, Anastas P. Tenev, Jan Zapal, and the participants of the 19th European Meeting on Game Theory (SING19) for their valuable comments and suggestions.

Funding Open access funding provided by Corvinus University of Budapest.

Declarations

Conflicts of Interest None.

References

Alonso, R., & Câmara, O. (2016). Bayesian persuasion with heterogeneous priors. *Journal of Economic Theory, 165*, 672–706.

Ambrus, A., & Egorov, G. (2017). Delegation and nonmonetary incentives. *Journal of Economic Theory, 171*, 101–135.

Augias, V., & Barreto, D. (2020). Persuading a wishful thinker. arXiv preprint arXiv:2011.13846.

Austen-Smith, D., & Banks, J. S. (2000). Cheap talk and burned money. *Journal of Economic Theory, 91*(1), 1–16.

Babichenko, Y., Talgam-Cohen, I., Xu, H., & Zabarnyi, K. (2022). Regret-minimizing Bayesian persuasion. *Games and Economic Behavior, 136*, 226–248.

Brooks, B., Frankel, A., & Kamenica, E. (2022). Information hierarchies. *Econometrica, 90*(5), 2187–2214.

Che, Y.-K., & Kartik, N. (2009). Opinions as incentives. *Journal of Political Economy, 117*(5), 815–860.

de Clippel, G., & Zhang, X. (2022). Non-Bayesian persuasion. *Journal of Political Economy, 130*(10), 2594–2642.

Dworczak, P., & Kolotilin, A. (2024). The persuasion duality. *Theoretical Economics, 19*(4), 1701–1755.

Eliaz, K., Spiegler, R., & Thysen, H. C. (2021). Persuasion with endogenous misspecified beliefs. *European Economic Review, 134*, Article 103712.

Eliaz, K., Spiegler, R., & Thysen, H. C. (2021). Strategic interpretations. *Journal of Economic Theory, 192*, Article 105192.
Enke, B., & Zimmermann, F. (2019). Correlation neglect in belief formation. *The Review of Economic Studies, 86*(1), 313–332.
Eyster, E., & Rabin, M. (2005). Cursed equilibrium. *Econometrica, 73*(5), 1623–1672.
Eyster, E., & Weizsäcker, G. (2011). Correlation neglect in financial decision-making, diw discussion papers, no. 1104.
Galperti, S. (2019). Persuasion: The art of changing worldviews. *American Economic Review, 109*(3), 996–1031.
Gentzkow, M., & Kamenica, E. (2016). Competition in persuasion. *The Review of Economic Studies, 84*(1), 300–322.
Hagmann, D., & Loewenstein, G. (2017). Persuasion with motivated beliefs. In *Opinion Dynamics & Collective Decisions Workshop*.
Herings, P.J.-J., Karos, D., & Kerman, T. T. (2024). Belief inducibility and informativeness. *Theory and Decision, 96*(4), 517–553.
Ichihashi, S. (2019). Limiting sender's information in bayesian persuasion. *Games and Economic Behavior, 117*, 276–288.
Kallir, I., & Sonsino, D. (2009). The neglect of correlation in allocation decisions. *Southern Economic Journal, 75*(4), 1045–1066.
Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review, 101*(6), 2590–2615.
Kartik, N. (2007). A note on cheap talk and burned money. *Journal of Economic Theory, 136*(1), 749–758.
Khemlani, S. S., Lotstein, M., & Johnson-Laird, P. N. (2015). Naive probability: Model-based estimates of unique events. *Cognitive Science, 39*(6), 1216–1258.
Kosterina, S. (2022). Persuasion with unknown beliefs. *Theoretical Economics, 17*(3), 1075–1107.
Levy, G., Barreda, I. M., & Razin, R. (2022). Persuasion with correlation neglect: a full manipulation result. *American Economic Review: Insights, 4*(1), 123–138.
Levy, G., & Razin, R. (2015). Correlation neglect, voting behavior, and information aggregation. *American Economic Review, 105*(4), 1634–1645.
Li, H. (2022). Transparency and policymaking with endogenous information provision. arXiv preprint arXiv:2204.08876.
Malamud, S., & Schrimpf, A. (2021). Persuasion by dimension reduction. arXiv preprint arXiv:2110.08884.
Ortoleva, P., & Snowberg, E. (2015). Overconfidence in political behavior. *American Economic Review, 105*(2), 504–535.
Rayo, L., & Segal, I. (2010). Optimal information disclosure. *Journal of Political Economy, 118*(5), 949–987.
Shimojo, S., & Ichikawa, S. (1989). Intuitive reasoning about probability: Theoretical and experimental analyses of the "problem of three prisoners''. *Cognition, 32*(1), 1–24.
Tamura, W. (2018). Bayesian persuasion with quadratic preferences. *Available at SSRN 1987877*.
Tsakas, E., Tsakas, N., & Xefteris, D. (2021). Resisting persuasion. *Economic Theory, 72*(3), 723–742.
Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293.
Wolfe, C. R., & Reyna, V. F. (2010). Semantic coherence and fallacies in estimating joint probabilities. *Journal of Behavioral Decision Making, 23*(2), 203–223.
Wyer, R. S., Jr. (1976). An investigation of the relations among probability estimates. *Organizational Behavior and Human Performance, 15*(1), 1–18.
Zhao, J., Shah, A., & Osherson, D. (2009). On the provenance of judgments of conditional probability. *Cognition, 113*(1), 26–36.