

Working Paper Series
(ISSN 2788-0443)

773

Changing Simplistic Worldviews

Maxim Senkov
Toygar T. Kerman

CERGE-EI
Prague, February 2024

**ISBN 978-80-7343-580-6 (Univerzita Karlova, Centrum pro ekonomický výzkum
a doktorské studium)**
ISBN 978-80-7344-707-6 (Národohospodářský ústav AV ČR, v. v. i.)

Changing Simplistic Worldviews*

Maxim Senkov[†] Toygar T. Kerman[‡]

January 19, 2024

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*: he ignores the dependency between the two dimensions of the state. We provide a characterization for the sender’s gain from persuasion both when the receiver is naive and when he is rational. 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; misspecified prior; correlation neglect

JEL Classification: D82, D83, D91

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 and Gentzkow \(2011\)](#) consider a prosecutor (sender) and a judge (receiver), both of whom are only interested in one dimension of the 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

*We would like to thank Ole Jann, Artyom Jelnov, Pavel Kocourek, Fedor Sandomirskiy, and Jan Zapal for their valuable comments and suggestions.

[†]European Research University, U Haldy 200/18, 700 30 Ostrava, Czech Republic. email: maxim.senkov@eruni.org.

[‡]Corresponding author. Institute of Economics, Corvinus University of Budapest, Fővám tér 8, 1093 Budapest, Hungary. email: toygar.kerman@uni-corvinus.hu. ORCID: 0000-0003-3038-3666. The author acknowledges funding by the Hungarian National Research, Development and Innovation Office, Project Number K-143276.

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. 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 evidence suggests that people find it challenging to work with joint distributions of random variables and tend to underestimate or ignore the correlation between state variables or observed signals (Kallir and Sonsino, 2009; Eyster and Weizsäcker, 2011; Enke and 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 and Gentzkow (2011) by (i) allowing for two-dimensional states of the world and (ii) considering a receiver who ignores the dependency of state variables (which leads to heterogeneous priors). We model the receiver’s disregard of state dependency as a misspecified prior belief; while the receiver knows the correct marginal distribution for each dimension, he assumes that they are independent. We characterize optimal disclosure both when the receiver is *rational* (i.e. does not ignore dependency) and *naive* (i.e. assumes that dimensions are independent), and analyze the welfare effects of the receiver’s simplistic worldview.

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

| | | | |
|-----|------|-----|------|
| | L | P | |
| U | 0.25 | 0.1 | 0.35 |
| S | 0.35 | 0.3 | 0.65 |
| | 0.6 | 0.4 | |

Table 1: Prior belief distribution.

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 $(0.65) \cdot 0 + (0.35) \cdot 1 = 0.35$ to the consultant and $(0.6) \cdot 1 + (0.4) \cdot 0 = 0.6$ 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 | 0.5 | 1 | 1 |
| t | 1 | 0.5 | 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 when it is sustainable and loss-making.¹ 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. The consultant's expected payoff of employing π is $(0.3) \cdot 1 + (0.35) \cdot 1 + 0.1 \cdot (0.5) + (0.25) \cdot 1 = 0.95$, improving upon the case of no disclosure. Similarly, the company's expected payoff from π is $(0.3) \cdot 1 + (0.35) \cdot 0 + (0.1) \cdot (0.5) + (0.25) \cdot 1 = 0.6$, 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 knows the marginal distributions $\Pr(S) = 0.65$ and $\Pr(P) = 0.4$, it assumes that $\Pr(S, P) = \Pr(S) \Pr(P) = (0.65) \cdot (0.4) = 0.26$.² The company's *simplistic worldview* regarding the prior distribution is given as follows.

¹Note that the optimal signal is not unique.

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

| | | | |
|-----|------|------|------|
| | L | P | |
| U | 0.21 | 0.14 | 0.35 |
| S | 0.39 | 0.26 | 0.65 |
| | 0.6 | 0.4 | |

Table 2: Misspecified prior belief distribution.

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 | 0.93 | 1 | 1 |
| t | 1 | 0.07 | 0 | 0 |

Indeed, when the project is unsustainable and profitable, the recommendation to continue is sent with probability 0.93, while it was sent with probability 0.5 in this state before. The consultant's expected payoff of employing $\hat{\pi}$ is $(0.3) \cdot 1 + (0.35) \cdot 1 + (0.1) \cdot (0.07) + (0.25) \cdot 1 = 0.907$, 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 $(0.3) \cdot 1 + (0.35) \cdot 0 + (0.1) \cdot (0.93) + (0.25) \cdot 1 = 0.643$, 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 (i) the sender's approach regarding disclosure in states with misaligned preferences and (ii) 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 characterization of the sender's gain from persuasion, and pin down the conditions under which the receiver benefits from having a simplistic worldview.

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

1.2 Related literature

Our paper extends [Kamenica and 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. 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, Barreda, and Razin \(2022\)](#)) is that our receiver ignores dependency between *states*, whereas most papers in the literature focus on the receiver neglecting correlation between *messages*.⁴

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.⁵

Our behavioral assumption regarding the receiver alters his perception of the joint belief distribution. Yet, the receiver has no ambiguity about the message he observes. [Eliaz, Spiegler, and Thyssen \(2021a\)](#) 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, Spiegler, and Thyssen \(2021b\)](#) 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.

Second, our paper contributes to the literature on Bayesian persuasion with heterogeneous priors. [Alonso and Câmara \(2016\)](#) assume that the sender and receiver have subjective prior beliefs and analyze its effect on the set of inducible posterior beliefs. [Galperti \(2019\)](#) focuses on a similar setup and considers receivers who might be reluctant to change their worldviews. In a related paper, [Kosterina \(2022\)](#) considers a sender who is ignorant about the receiver’s prior and characterizes optimal information structures. In contrast to these studies, our model posits that the sender is aware of the accurate

⁴[Eyster and Rabin \(2005\)](#) introduce *cursed equilibrium* which assumes that players underestimate the correlation between other players’ actions and information.

⁵This also relates to studies that investigate how a receiver can strategize to increase the informativeness of the signal ([Austen-Smith and Banks, 2000](#); [Kartik, 2007](#); [Ambrus and Egorov, 2017](#); [Tsakas, Tsakas, and Xefteris, 2021](#)) and how different kinds of biases might affect persuasion ([Hagmann and Loewenstein, 2017](#); [Augias and Barreto, 2020](#); [de Clippel and Zhang, 2022](#)).

prior distribution while the receiver is not, and this discrepancy is common knowledge. Moreover, we contrast the receiver’s payoff to the case in which he has the correct prior belief and study the implications of incorrect prior for welfare.

Third, our study relates to the literature on persuasion with multidimensional state variables. [Rayo and Segal \(2010\)](#) consider a model of information disclosure and assume that the agent is rational, 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, Talgam-Cohen, Xu, and Zabarnyi \(2022\)](#) 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.

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 $\omega_{k\ell} \equiv (\sigma_k, \rho_\ell) \in \Omega$.

A *signal* consists of a finite set of *messages* M and a family of distributions $\{\pi(\cdot|\omega)\}_{\omega \in \Omega}$ over M . That is, $\pi : \Omega \rightarrow \Delta(M)$ maps each state of the world to a probability distribution over M . Denote the set of all signals by Π .

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 $\mu(\rho_\ell) \equiv \Pr(\rho_\ell)$. The distribution μ is presented in [Table 3](#) with marginal distributions given on the sides.

| | | | |
|------------------------|------------|------------|-----------------|
| (Σ, \mathbb{P}) | ρ_0 | ρ_1 | $\mu(\sigma_0)$ |
| σ_0 | μ_{00} | μ_{01} | |
| σ_1 | μ_{10} | μ_{11} | |

$\mu(\rho_0)$

Table 3: Common prior μ .

Given $\mu \in \Delta(\Omega)$ and $\pi \in \Pi$, a message $m \in M^\pi$ 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

$$\mu^m(\rho_\ell) = \frac{\mu_{k\ell}\pi(m|\omega_{k\ell}) + \mu_{\ell\ell}\pi(m|\omega_{\ell\ell})}{\mu_{k\ell}\pi(m|\omega_{k\ell}) + \mu_{\ell\ell}\pi(m|\omega_{\ell\ell}) + \mu_{\ell k}\pi(m|\omega_{\ell k}) + \mu_{k k}\pi(m|\omega_{k k})}$$

and $\mu^m(\sigma_k)$ is given analogously.⁶ We denote the posterior belief for the joint distribution upon observing message m by $\mu^m(\sigma_k, \rho_\ell)$.

The set of actions that are available to the receiver is given by $A = \{0, 1\}$. 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 $u : A \times R \rightarrow \{0, 1\}$ 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 $v : A \times \Sigma \rightarrow \{0, 1\}$ is given by

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

We assume that the the default action of the receiver is 0, i.e. $\mu(\rho_1) < \mu(\rho_0)$. Given $\pi \in \Pi$, the *choice rule* $\alpha^\pi : M \rightarrow A$ of the receiver is given by

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

That is, the receiver switches his action when he is indifferent upon observing a message.

⁶We drop π when denoting $\mu^m(\sigma_k)$ and $\mu^m(\rho_\ell)$ to ease notation.

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.

First, note that since both the receiver's action set and the payoff-relevant state are binary, we can restrict attention to *direct* signals without loss of generality.⁷ Therefore, we hereon assume that $M = \{0, 1\}$ and that Π is the set of all direct signals. 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 $k, \ell \in \{0, 1\}$.

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 $\mu^1(\rho_1) \geq 1/2$. We formalize the obedience constraint in the following lemma.

Lemma 1. *Given $\pi \in \Pi$, the receiver chooses action 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 obedience 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:

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

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

Proposition 1. *An optimal signal when the receiver is rational 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

⁷A signal is *direct* if it holds that (i) $M \subseteq A$ and (ii) for all $a \in M$, $\alpha^\pi(a) = a$.

the state is (σ_0, ρ_0) the sender has an incentive to *decrease* the frequency of message 1. Note that both increasing π_{11} and decreasing π_{00} *relaxes* the obedience constraint in (2).

Second, 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 obedience constraint since the sender's and receiver's preferences are misaligned. Therefore, the sender chooses π_{01} and π_{10} such that the obedience constraint binds. However, if the prior for (σ_1, ρ_1) is (weakly) higher than the prior for (σ_1, ρ_0) (i.e. $\mu_{11} \geq \mu_{10}$), then the effect of relaxing the constraint is stronger than the effect of tightening the constraint. In this case, since $\pi_{01}^* = 0$ and $\pi_{10}^* = 1$, the sender can implement her desired outcome with certainty and obtain a payoff of 1.

3 Optimal signal when the receiver is naive

From hereon 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. $\hat{\mu}_{k\ell} = \mu_k^\sigma \mu_\ell^\rho$. We present the misspecified prior in the following table with marginal distributions written on the sides.

| | | | |
|---------------|---------------------------|---------------------------|-----------------|
| (Σ, P) | ρ_0 | ρ_1 | |
| σ_0 | $\mu_0^\sigma \mu_0^\rho$ | $\mu_0^\sigma \mu_1^\rho$ | $\mu(\sigma_0)$ |
| σ_1 | $\mu_1^\sigma \mu_0^\rho$ | $\mu_1^\sigma \mu_1^\rho$ | |
| | $\mu(\rho_0)$ | | |

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

We call such a misspecified prior $\hat{\mu}$ a *simplistic worldview*. Lemma 2 below demonstrates that the simplistic worldview relates to the correct prior in a systematic way.

Lemma 2. *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 2 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 demonstrates the misspecified prior $\hat{\mu}$ in terms of the correct prior μ .

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

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

The naive receiver observes the signal chosen by the sender, who knows that the receiver is naive. Moreover, he 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

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

While the sender calculates her expected utility according to μ , she knows that the naive receiver updates according to $\hat{\mu}$. Therefore, the sender's optimization problem in (2) becomes

$$\begin{aligned} \max_{\pi \in \Pi} \quad & \mu_{11}\pi_{11} + \mu_{10}\pi_{10} - \mu_{01}\pi_{01} - \mu_{00}\pi_{00} \\ \text{s.t.} \quad & \hat{\mu}_{00}\pi_{00} + \hat{\mu}_{10}\pi_{10} \leq \hat{\mu}_{01}\pi_{01} + \hat{\mu}_{11}\pi_{11}. \end{aligned} \quad (4)$$

Let $\hat{\pi}^*$ denote the optimal signal when the receiver is naive. Note that in (4), similar to (2), the sender's optimal choice in states (σ_0, ρ_0) and (σ_1, ρ_1) relaxes the obedience constraint and thus $\hat{\pi}_{00}^* = 0$ and $\hat{\pi}_{11}^* = 1$, which allows us to simplify (4).

Lemma 3. *The sender's simplified problem is given by*

$$\begin{aligned} \max_{\pi \in \Pi} \quad & \mu_{10}\pi_{10} - \mu_{01}\pi_{01} \\ \text{s.t.} \quad & \hat{\mu}_{10}\pi_{10} - \hat{\mu}_{01}\pi_{01} \leq \hat{\mu}_{11}. \end{aligned} \quad (5)$$

We now provide the full characterization of the optimal signal for a naive receiver.

Proposition 2. *An optimal signal when the receiver is naive 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. 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 obedience 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, however, 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 obedience constraint to a different extent. Hence, the sender chooses the option that tightens the obedience constraint less as to optimally persuade the receiver.

4 Welfare implications of a simplistic worldview

In this section, we proceed beyond characterizing the optimal signal and study a number of general properties of the sender's problem (4) 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 $c > 0$.*

(i) *For any $\pi \in \Pi$ with $\pi_{00} = 0$ and $\pi_{11} = 1$ it holds that $\hat{\mu}^1(\rho_1) > \mu^1(\rho_1)$. 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) *For any $\pi \in \Pi$ it holds that $(\hat{\mu}_{11}/\hat{\mu}_{10}) + (\hat{\mu}_{01}/\hat{\mu}_{10})\pi_{01} > (\mu_{11}/\mu_{10}) + (\mu_{01}/\mu_{10})\pi_{01}$. That is, the obedience constraint in (5) 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 2 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.

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 obedience constraint in (5) is more relaxed relative to the benchmark (2). It is easy to see that the converse of the above arguments 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 obedience 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 $c > 0$. A comparison of expected utilities yields the following:*

(i) *The receiver’s simplistic worldview harms him and benefits the sender.*

(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 get her preferred outcome with certainty 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 get her preferred outcome with certainty 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 2 if $c > 0$ then the naive receiver perceives the aligned states as more likely than the rational receiver. Therefore, Lemma 4 implies that it is easier for the sender to persuade the receiver. Hence, *naivete makes the receiver more manipulable and is detrimental*. 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*.

5 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.

Appendix A Proofs

Proof of Lemma 1. For the obedience 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 obedience constraint. \square

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 obedience constraint for 0 is satisfied. That is,

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

\square

Proof of Proposition 1. 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 obedience constraint in (2) and maximizes the sender’s expected utility since it provides the upper bound 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 obedience constraint. Hence, substituting $\pi_{11}^* = 1$ and $\pi_{00}^* = 0$ into the sender's optimization problem and simplifying we obtain

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

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). \square

Proof of Lemma 2. Let $\hat{\mu}_{kk} - \mu_{kk} = c$ and assume that $k \neq \ell$. Then,

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

\square

Proof of Lemma 3. As one can observe from (4), the sender's payoff is increasing in π_{11} , which relaxes the obedience constraint. Thus, $\hat{\pi}_{11}^* = 1$. Similar reasoning leads to $\hat{\pi}_{00}^* = 0$. Substituting $\hat{\pi}_{00}^* = 0, \hat{\pi}_{11}^* = 1$ into (4) yields the result. \square

Proof of Proposition 2. The sender's problem is given by

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

As the problem is a linear program, the solution lies either on the obedience 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 obedience 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 Figure 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 (5) has this particular form if and only if the slope of the objective function is steeper than the slope of the obedience 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 (5) 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

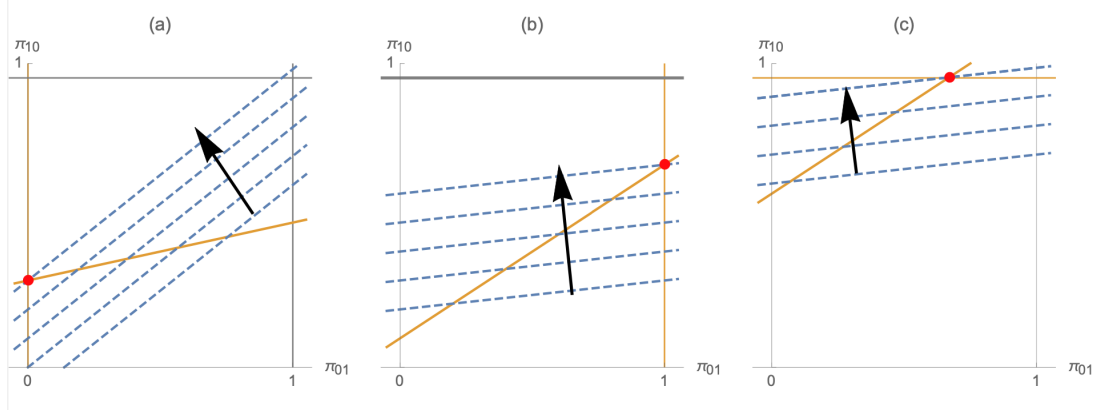


Figure 1: Three possible forms of solutions to the problem (5). Indifference curves are in blue and dashed. Active constraints are in orange.

is flatter than the slope of the obedience 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. $\hat{\pi}_{01}^* = (\hat{\mu}_{10} - \hat{\mu}_{11})/\hat{\mu}_{01}$. The solution to (5) 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 (5). \square

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 2, it follows that

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

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 (5) for the naive receiver. Rewriting the constraints in (2) and (5), 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 2 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 (5).

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 (5). \square

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,

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

As the form of optimal signal in Propositions 1 and 2 depends on $\text{sign}(\mu_{11} - \mu_{10})$, we consider four cases based on $\text{sign}(\mu_{11} - \mu_{10})$ and $\text{sign}(c)$. First, assume $\mu_{11} \geq \mu_{10}$. If $c \geq 0$, then from Proposition 1 and 2 and Lemma 2, $(\pi_{00}^*, \pi_{01}^*, \pi_{10}^*, \pi_{11}^*) = (\hat{\pi}_{00}^*, \hat{\pi}_{01}^*, \hat{\pi}_{10}^*, \hat{\pi}_{11}^*) = (0, 0, 1, 1)$. 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

$$\hat{V}^* \geq V^* \text{ if and only if } c \geq 0. \quad (7)$$

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

$$\begin{aligned} \nu &\equiv \hat{V}^* - V^* = \mu_{10}(\hat{\pi}_{10}^* - \pi_{10}^*) + \mu_{01}(\pi_{01}^* - \hat{\pi}_{01}^*) \\ &= \mu_{10}(\hat{\pi}_{10}^* - \gamma) + \mu_{01}\left(\frac{\mu_{10}}{\mu_{01}}\gamma - \frac{\mu_{11}}{\mu_{01}} - \hat{\pi}_{01}^*\right) \\ &= \mu_{10}\hat{\pi}_{10}^* - \mu_{01}\hat{\pi}_{01}^* - \mu_{11}. \end{aligned} \quad (8)$$

Plugging the optimal signal for case (a) from Proposition 2 in (8) yields

$$\begin{aligned}\nu &= \mu_{10} \frac{\hat{\mu}_{11}}{\hat{\mu}_{10}} - \mu_{11} = \mu_{10} \frac{\mu_{11} + c}{\hat{\mu}_{10}} - \mu_{11} \\ &= \frac{\mu_{10}\mu_{11} + \mu_{11}c - \mu_{10}\mu_{11} + \mu_{10}c}{\hat{\mu}_{10}} = \frac{c(\mu_{11} + \mu_{10})}{\hat{\mu}_{10}}.\end{aligned}\quad (9)$$

Similarly, plugging the optimal signal for case (b) from Proposition 2 in (8) yields

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

Finally, plugging the optimal signal for case (c) from Proposition 2 in (8) yields

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

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 2, this can be rewritten as $\mu(\rho_1) - \mu_{10} \geq -c$. Since we also have $\mu_{01} \geq c$ by Lemma 2, it holds that

$$\mu(\rho_1) - \mu_{10} + \mu_{01} \geq 0. \quad (12)$$

By (9), (10), (11), and (12) it follows that,

$$\hat{V}^* \geq V^* \text{ if and only if } c \geq 0. \quad (13)$$

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:

$$\begin{aligned}& \underbrace{\mu_{11}\pi_{11}^* + \mu_{00}(1 - \pi_{00}^*) + \mu_{10}\pi_{10}^* + \mu_{01}(1 - \pi_{01}^*)}_{\text{the sender's expected payoff}} \\ & \quad + \underbrace{\mu_{11}\pi_{11}^* + \mu_{00}(1 - \pi_{00}^*) + \mu_{10}(1 - \pi_{10}^*) + \mu_{01}\pi_{01}^*}_{\text{the receiver's expected payoff}} \\ &= 2(\mu_{11}\pi_{11}^* + \mu_{00}(1 - \pi_{00}^*)) + \mu_{10} + \mu_{01} \\ &= 2(\mu_{11} + \mu_{00}) + \mu_{10} + \mu_{01},\end{aligned}\quad (14)$$

where both use the correct (sophisticated) prior belief for evaluating the receiver's expected utility when he is naive.

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 (14) 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 (7) and (13), it follows that

$$\hat{U}^* \leq U^* \text{ if and only if } c \geq 0, \quad (15)$$

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 $c > 0$, from Lemma 4, the receiver's naivete relaxes the obedience constraint. Thus, the difference between the expected payoffs in (7), (13), and (15) 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 obedience constraint. Thus, the difference between the expected payoffs in (7), (13), and (15) 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 $\hat{\mu}_{11} < \hat{\mu}_{10}$. \square

References

- Alonso, R. and O. Câmara (2016). Bayesian persuasion with heterogeneous priors. *Journal of Economic Theory* 165, 672–706.
- Ambrus, A. and G. Egorov (2017). Delegation and nonmonetary incentives. *Journal of Economic Theory* 171, 101–135.
- Augias, V. and D. Barreto (2020). Persuading a wishful thinker. *arXiv preprint arXiv:2011.13846*.
- Austen-Smith, D. and J. S. Banks (2000). Cheap talk and burned money. *Journal of Economic Theory* 91(1), 1–16.
- Babichenko, Y., I. Talgam-Cohen, H. Xu, and K. Zabarnyi (2022). Regret-minimizing Bayesian persuasion. *Games and Economic Behavior* 136, 226–248.
- de Clippel, G. and X. Zhang (2022). Non-Bayesian persuasion. *Journal of Political Economy* 130(10), 2594–2642.
- Dworczak, P. and A. Kolotilin (2019). The persuasion duality. *arXiv preprint arXiv:1910.11392*.
- Eliaz, K., R. Spiegler, and H. C. Thysen (2021a). Persuasion with endogenous misspecified beliefs. *European Economic Review* 134, 103712.
- Eliaz, K., R. Spiegler, and H. C. Thysen (2021b). Strategic interpretations. *Journal of Economic Theory* 192, 105192.

- Enke, B. and F. Zimmermann (2019). Correlation neglect in belief formation. *The Review of Economic Studies* 86(1), 313–332.
- Eyster, E. and M. Rabin (2005). Cursed equilibrium. *Econometrica* 73(5), 1623–1672.
- Eyster, E. and G. Weizsäcker (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.
- Hagmann, D. and G. Loewenstein (2017). Persuasion with motivated beliefs. In *Opinion Dynamics & Collective Decisions Workshop*.
- Kallir, I. and D. Sonsino (2009). The neglect of correlation in allocation decisions. *Southern Economic Journal* 75(4), 1045–1066.
- Kamenica, E. and M. Gentzkow (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.
- Kosterina, S. (2022). Persuasion with unknown beliefs. *Theoretical Economics* 17(3), 1075–1107.
- Levy, G., I. M. d. Barreda, and R. Razin (2022). Persuasion with correlation neglect: a full manipulation result. *American Economic Review: Insights* 4(1), 123–138.
- Levy, G. and R. Razin (2015). Correlation neglect, voting behavior, and information aggregation. *American Economic Review* 105(4), 1634–1645.
- Malamud, S. and A. Schrimpf (2021). Persuasion by dimension reduction. *arXiv preprint arXiv:2110.08884*.
- Ortoleva, P. and E. Snowberg (2015). Overconfidence in political behavior. *American Economic Review* 105(2), 504–535.
- Rayo, L. and I. Segal (2010). Optimal information disclosure. *Journal of political Economy* 118(5), 949–987.
- Tamura, W. (2018). Bayesian persuasion with quadratic preferences. *Available at SSRN 1987877*.
- Tsakas, E., N. Tsakas, and D. Xefteris (2021). Resisting persuasion. *Economic Theory* 72(3), 723–742.

Abstrakt

Studujeme bayesovský model přesvědčování s dvourozměrnými stavy světa, v němž mají odesílatel (ona) a příjemce (on) různá předchozí přesvědčení a zajímají se o různé dimenze. Příjemce je naivní agent, který má zjednodušený pohled na svět: ignoruje závislost mezi dvěma dimenzemi stavu. Poskytujeme charakteristiku zisku odesílatele z přesvědčování jak v případě, že je příjemce naivní, tak v případě, že je racionální. Ukazujeme, že příjemce má prospěch ze zjednodušeného pohledu na svět tehdy a jen tehdy, když díky němu vnímá stavy, v nichž je jeho zájem v souladu s odesílatelem, jako méně pravděpodobné.

Working Paper Series
ISSN 2788-0443

Individual researchers, as well as the on-line version of the CERGE-EI Working Papers (including their dissemination) were supported from institutional support RVO 67985998 from Economics Institute of the CAS, v. v. i.

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

(c) Maxim Senkov, Toygar T. Kerman, 2024

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical or photocopying, recording, or otherwise without the prior permission of the publisher.

Published by
Charles University, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute of the CAS, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Phone: + 420 224 005 153
Email: office@cerge-ei.cz
Web: <https://www.cerge-ei.cz/>

Editor: Byeongju Jeong

The paper is available online at <https://www.cerge-ei.cz/working-papers/>.

ISBN 978-80-7343-580-6 (Univerzita Karlova, Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-707-6 (Národohospodářský ústav AV ČR, v. v. i.)