Persuasion with Correlation Neglect: A Full Manipulation Result

By Gilat Levy, Inés Moreno de Barreda, Ronny Razin*

We consider an information design problem in which a sender tries to persuade a receiver that has "correlation neglect", i.e. fails to understand that signals might be correlated. We show that a sender with unlimited number of signals can fully manipulate the receiver. Specifically, the sender can induce the receiver to hold any state-dependent posterior she wishes to. If the sender only wishes to induce a state-independent posterior, she can use fully correlated signals, but generally she needs to design more involved correlation structures.

In this day and age we are constantly exposed to many different information sources, including a myriad of social media platforms. The growth in the amount of information we are exposed to and its complexity may open up opportunities for those who try to manipulate our beliefs. Indeed, in the last few years, more and more instances of orchestrated attempts to manipulate information using social media have been exposed, and companies like Facebook are under increasing pressure to take action against what they call "coordinated inauthentic behaviour" on its platforms.[1]

While information sources may be correlated, individuals may be unaware that this is the case. A recent empirical literature documents that individuals exhibit "correlation neglect". Formally, this means that they assume that information sources are (conditionally) independent. Ortoleva and Snowberg (2015) documents how correlation neglect shapes political views. Eyster and Weizsäker (2011), Kallir and Sonsino (2009) and Enke and Zimmermann (2019) provide experimental evidence for such behaviours.[2]

In this paper we investigate how coordination across information sources, when unknown to the individual, can be used strategically to affect her beliefs. We analyse a model of persuasion with one sender and one receiver. We use an information design framework in which a sender can design and commit to a joint information structure with m signals as a function of the state of the world. The receiver, who attempts to learn the state of the world, observes the realisations of these m signals. The receiver understands the marginal distribution of each signal but she updates her beliefs assuming that

* Levy: London School of Economics (e-mail: g.levy1@lse.ac.uk); Moreno de Barreda: University of Oxford (e-mail: ines.morenodebarreda@economics.ox.ac.uk); Razin: London School of Economics (e-mail: r.razin@lse.ac.uk). This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 681579. The results in this paper appeared in a previous version entitled: "Persuasion with Correlation Neglect" (Levy, Moreno de Barreda and Razin, 2020), which was split into two working papers, the current one and another entitled "Feasible Joint Posterior Distributions: A Graphical Approach" (Levy, Moreno de Barreda and Razin, 2021*a*).

[1] See for example https://about.fb.com/news/2020/10/removing-coordinated-inauthentic-behavior-mexico-iran-myanmar/.

[2] See also De Marzo, Vayanos and Zwiebel (2003), Golub and Jackson (2012), Gagnon-Bartsch and Rabin (2016), Glaeser and Sunstein (2009) and Levy and Razin (2015*a*;*b*).

the signals are all conditionally independent.

Our main result shows that the sender can achieve full manipulation as the number of signals she can use grows large. That is, in the limit she can induce the receiver to hold any state-dependent belief. With continuity assumptions on utilities, this implies that the sender can approximate her first-best state-dependent utility.

To appreciate this result consider the following simple example, which we will revisit later on. The state of the world is binary, $\Omega = \{0, 1\}$, with uniform prior. We identify a posterior with the probability it allocates to the state $\omega = 1$. The sender wants to confuse the receiver in both states, i.e. she would like the receiver to have a posterior higher or equal than a threshold $q > 1/2$ in state 0, and lower or equal than $1 - q$ in state 1.

If the sender could only use one signal, she can achieve her goal with a probability of at most $1-q$. She can do that by using a binary signal with accuracy q; the signal s takes two realisations 0 or 1 with $Pr(s = 1 \mid \omega = 1) = Pr(s = 0 \mid \omega = 0) = q$. Such signal induces two posteriors $\mu_1 = q$ and $\mu_0 = 1 - q$. In particular, posterior μ_1 is induced with probability q in state 1 and probability $1 - q$ in state 0 and posterior μ_0 is induced with the complementary probabilities.

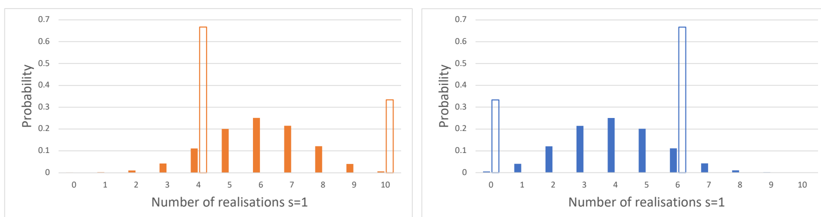
However, we show that as the number of signals she can use grows large, the sender is able to achieve her goal with probability approaching one. The intuition comes from the combination of two observations. First, correlation neglect creates an amplification effect. As the receiver thinks signals are independent, she expects the realisations to behave as a multinomial distribution. As the number of signals increases, the law of large numbers implies that by using a set of realisations that deviate slightly from the expectation, the sender can induce extreme posteriors. This is the case even if the signals are relatively uninformative.

The second observation is that in order to achieve full manipulation, the sender needs to exploit negative correlation. Coming back to the example, for a realisation to induce a posterior $\mu_0 < 1/2$, such a realisation has to be sent with a higher probability when the state is 0 than when the state is 1. Hence, if μ_0 is induced with high probability in state 1, then it has to be induced with even higher probability in state 0. Therefore, if the signals were fully positively correlated, it becomes impossible to induce state-dependent beliefs that are contrarian to the true state with high probability.

While the sender needs to use negative correlation, there are limits to the amount of negative correlation of posteriors that can be induced with a signal structure (Levy, Moreno de Barreda and Razin, 2021*a*). For example, the sender cannot simultaneously send two signals inducing $\mu_0 = 0$ and $\mu_1 = 1$ as each such posterior can only be sent in its respective state. However, if instead she uses signals that are almost uninformative, with posteriors $\mu_0 = 1/2 - \varepsilon$ and $\mu_1 = 1/2 + \varepsilon$ for a small ε, she can simultaneously send μ_0 and μ_1 with a very high probability.

The combination of these two observations implies that as the number of signals grows large, the sender can use relatively uninformative signals and correlate them in the right manner, to generate any desired posterior at any state with high probability.

Figure 1 illustrates the intuition for the example above. Suppose that the posterior threshold for the receiver is $q = 2/3$ and that the sender uses 10 binary signals with accuracy 0.6. The horizontal axis represents the number of signals with realisation $s = 1$, and the vertical axis represents the probability of such events. The solid bars correspond to the binomial distribution that the receiver thinks she is facing in each of the two states of the world. Instead, the sender correlates the realisations of the signals in accordance with the blank bars: In state $\omega = 1$, with probability $1/3$ all of the signals have realisation $s = 1$ and with probability $2/3$ only 4 of them have realisation $s = 1$. In this latter case, the receiver's posterior is equal to the desired $1 - q = 1/3$, since the binomial probability of observing exactly 4 realisations $s = 1$ is twice as likely in state $\omega = 0$ than in state $\omega = 1$. The opposite structure is used in state $\omega = 0$. The sender then induces the desired posteriors with probability $2/3$ rather than the initial $1/3$ that was achieved with a single signal. And as the sender adds more signals to her arsenal, she approaches full manipulation.



Panel A: State $\omega = 1$ Panel B: State $\omega = 0$

Figure 1. Real and perceived information structures conditional on the state.

Our paper relates to the growing literature on persuasion, stemming from Aumann and Maschler (1995) and Kamenica and Gentzkow (2011). A recent literature analyses persuasion in the presence of behavioural biases of the receiver. Our analysis should be viewed as a worst-case scenario; we allow the sender to fully commit to an information structure, the receiver does not perceive any level of correlation, and the sender has potentially a large number of signals at her disposal.[3] A similar worst-case scenario approach is taken in Eliaz, Spiegler and Weiss (Forthcoming). Alonso and Câmara (2016) analyse persuasion when the receiver has a wrong prior. In an example similar in spirit to the above they show, in contrast to our result, that full manipulation cannot be achieved.

Our analysis relates more generally to the literature on misspecified models. In other recent contributions, Eliaz and Spiegler (2020), Schumacher and Thysen (2018), and Ellis, Piccione and Zhang (Forthcoming) characterise how individuals with misperceived models can be strategically manipulated.

I. The model and preliminary results

There is a finite state space Ω, $\#\Omega = n$, with a full support prior probability distribution $p \in \Delta(\Omega)$ which is common knowledge. The sender designs a finite information struc-

[3]In Levy, Moreno de Barreda and Razin (2021*b*) we consider a related model that is applied to political competition.

ture that consists of $m \in \{1, 2, ...\}$ distinct signals. The receiver observes the m signal realisations and chooses an action a out of a compact set $A \subseteq \mathbb{R}$. Given action a and state ω, the receiver gets utility $u(a, \omega)$ and the sender gets utility $v(a, \omega)$.

A. *Information structures*

A finite information structure with m signals is defined by $\{\boldsymbol{S}, \{q(\cdot \mid \omega)\}_{\omega\in\Omega}\}$ where $\boldsymbol{S} = \prod_{i=1}^{m} S_i$ is the Cartesian product of the support of the individual signals and $q(\cdot \mid \omega) \in \Delta(\boldsymbol{S})$ is the joint probability distribution conditional on $\omega \in \Omega$. We assume S_i is finite for all $i \in \{1, ..., m\}$. Let $\{S_i, \{q_i(\cdot \mid \omega)\}_{\omega\in\Omega}\}$ denote the marginal information structure for signal i derived from $\{\boldsymbol{S}, \{q(\cdot \mid \omega)\}_{\omega\in\Omega}\}$, and let $q_i(\cdot)$ denote the marginal unconditional probability distribution.

As in the Bayesian Persuasion literature it will be more convenient to abstract from signal structures and work directly with distributions over posteriors. Given a realisation $s_i \in S_i$ of signal i with $q_i(s_i) > 0$, define the posterior induced by s_i to be,

$$\mu_{s_i}(\omega) = \frac{p(\omega)q_i(s_i \mid \omega)}{q_i(s_i)} \qquad \text{for all } \omega \in \Omega. \tag{1}$$

The joint distribution of the signals $\{q(\cdot \mid \omega)\}_{\omega\in\Omega}$ induces a family of conditional distributions, $\{\tau(\cdot \mid \omega)\}_{\omega\in\Omega} \subset \Delta(\Delta(\Omega)^m)$, over vectors of posteriors, such that[4]

$$\tau((\mu_{s_1}, ..., \mu_{s_m}) \mid \omega) = q((s_1, ..., s_m) \mid \omega) \quad \text{ for all } \omega \in \Omega. \tag{2}$$

Given the joint conditional distributions $\{\tau(\cdot \mid \omega)\}_{\omega\in\Omega}$, we denote by $\{\tau_i(\cdot \mid \omega)\}_{\omega\in\Omega} \subset \Delta(\Delta(\Omega))$ the set of marginal conditional distributions over posteriors corresponding to the $i'th$ signal and $\tau_i(\cdot) = \sum_{\omega\in\Omega} p(\omega)\tau_i(\cdot \mid \omega) \in \Delta(\Delta(\Omega))$ the corresponding marginal unconditional distribution.

Note that equations (1) and (2) impose conditions on the distributions over posteriors that might be generated from a signal structure. The following lemma characterises the families of joint conditional distributions that can be induced by m signals:[5]

LEMMA 1: *The set of joint conditional distributions $\{\tau(\cdot \mid \omega)\}_{\omega\in\Omega}$ is inducible by m signals if and only if for any $\mu \in \Delta(\Omega)$ and any $i \in \{1, ..., m\}$ such that $\tau_i(\mu) > 0$,*

$$\textit{(Conditional Constraints)} \qquad \mu(\omega) = \frac{p(\omega)\tau_i(\mu \mid \omega)}{\tau_i(\mu)} \quad \textit{for all } \omega \in \Omega. \tag{3}$$

Intuitively, each posterior needs to coincide with what a Bayesian updater would generate from the respective signal.[6] By Lemma 1, we can then abstract from explicit signal

[4]Here, wlog, we are implicitly assuming that for every signal i, distinct realisations s_i, s_i' lead to different posteriors $\mu_{s_i} \neq \mu_{s_i'}$. If for signal i there are $s_i \neq s_i' \in S_i$ such that $\mu_{s_i} = \mu_{s_i'}$, then we replace these two realisations by a new one s_i'' with $q_i(s_i'' \mid \omega) = q_i(s_i \mid \omega) + q_i(s_i' \mid \omega)$.

[5]The proof is trivial and therefore omitted. Similar results are in Gutmann et al. (1991), Arieli et al. (Forthcoming), Ziegler (2020) and Arieli, Babichenko and Sandomirskiy (2020).

[6]This condition is related to the *obedience* condition in Bergemann and Morris (2016). In their setup, the sender

structures and work with conditional distributions over posteriors that satisfy the conditional constraints.

B. *Correlation neglect heuristic*

Given any set of joint conditional distributions $\{\tau(\cdot \mid \omega)\}_{\omega\in\Omega}$, we assume that the receiver understands the marginal distributions $\{\tau_i(\cdot \mid \omega)\}_{\omega\in\Omega}$. However, she believes that for any vector of posteriors $\boldsymbol{\mu} = (\mu_1, ..., \mu_m)$, the conditional joint distribution is given by $\prod_{i=1}^{m} \tau_i(\mu_i \mid \omega)$. We denote by $\mu^{CN}(\boldsymbol{\mu}) \in \Delta(\Omega)$ the posterior that a receiver with correlation neglect forms when observing $\boldsymbol{\mu}$. The following characterises $\mu^{CN}(\boldsymbol{\mu})$.[7]

LEMMA 2: *Given a prior p and a vector of posteriors $\boldsymbol{\mu}$, the posterior belief of a receiver with correlation neglect is:*

$$\mu^{CN}(\boldsymbol{\mu})(\omega) = \frac{\prod_{i=1}^{m} \mu_i(\omega) \Big/ p(\omega)^{m-1}}{\sum_{\upsilon\in\Omega} \prod_{i=1}^{m} \mu_i(\upsilon) \Big/ p(\upsilon)^{m-1}}.$$

The formula in Lemma 2 is inherited from the multiplicative form of the joint distribution of independent signals. Note that μ^{CN} is symmetric in $\boldsymbol{\mu}$, and depends only on the realised vector of posteriors and not on the actual distribution over vectors of posteriors τ.

II. Complete manipulation

In this Section we first show how the sender can completely manipulate the beliefs of the receiver. We then make some mild assumptions on the utilities $u(a, \omega)$ and $v(a, \omega)$ to show that this manipulation implies that the sender can approach her first-best.

Let τ^m denote a signal structure with m signals, and recall that the sender designs τ^m subject to the conditional constraints and the receiver updating her beliefs according to the correlation neglect heuristic. Let $(\rho^\omega)_{\omega\in\Omega}$ denote a vector of state-dependent posteriors and let $|\cdot|$ denote the Euclidean distance in $\mathbb{R}^n$ ($\#\Omega = n$). Our main result shows that when the sender has a large number of signals in her arsenal, she is able to induce in state ω posteriors that are arbitrarily close to ρ^ω, with probability approaching one.

THEOREM 1: *For any $(\rho^\omega)_{\omega\in\Omega} \in \Delta(\Omega)^n$, there exists a sequence of signal structures $\{\tau^m\}_{m\in\mathbb{N}}$, such that for any $\omega \in \Omega$, and any $\epsilon > 0$,*

$$\lim_{m\to\infty} \tau^m(\{\boldsymbol{\mu} \text{ s.t. } \tau^m(\boldsymbol{\mu}) > 0 \mid |\mu^{CN}(\boldsymbol{\mu}) - \rho^\omega| < \epsilon\} \mid \omega) = 1.$$

recommends an action to the receiver and the recommendation should be consistent with what a rational Bayesian updater would do upon receiving such a recommendation.

[7]This result exists in Sobel (2014), Proposition 5. The multiplicative form also arises in Morris (1977) and Bordley (1982).

In the proof we use the fact that the receiver's perception of conditional independence allows the sender to make use of the law of large numbers while exploiting different forms of correlation. Specifically, we highlight the combination of two features. First, the correlation neglect heuristic creates an amplification effect. For example, repeating a posterior that is removed from the prior in a particular direction, amplifies the belief of the receiver in that direction. Second, by introducing negative correlation of posteriors, the sender can increase the conditional probability with which she can generate correlation neglect beliefs that go against the state. While in general the sender cannot negatively correlate posteriors with a high probability because of the conditional constraints, these constraints are relaxed as signals become more uninformative. As the receiver is very sensitive to small deviations from an uninformative signal, the sender can then expertly correlate such signals to approach any desired posterior at any state with probability close to one.

To see this more clearly, consider the following Example, that we discussed in the introduction:

EXAMPLE 1 (An "Evil" Sender): *Consider a binary state space* $\Omega = \{0, 1\}$ *with uniform prior. We identify a posterior with the probability it allocates to the state* $\omega = 1$. *The receiver has three possible actions* $A = \{L, M, H\}$. *Her optimal action is H if her posterior is on or above* $2/3$, *L if her posterior is on or below* $1/3$ *and M otherwise. The sender would like to induce action H in state 0 and action L in state 1. In other words, the sender wishes to confuse the receiver.*

If the sender has only one signal, her optimal information structure would generate posteriors $\mu_0 = 1/3$ and $\mu_1 = 2/3$ with the following conditional distributions: $\tau(\mu_0 \mid 1) = 1/3$, $\tau(\mu_1 \mid 1) = 2/3$, $\tau(\mu_0 \mid 0) = 2/3$ and $\tau(\mu_1 \mid 0) = 1/3$. She is then able to fool the receiver one third of the time.

If the sender can use more than one signal, she could use the amplification effect to generate the same posteriors with less informative signals. This will allow her to fool the receiver with higher probability. To see this take the simple example of fully correlated homogeneous signals. Given a posterior $\mu \in \Delta(\Omega)$, let $\mu_m^{FC}(\mu) \equiv \mu^{CN}(\mu, \mu, ..., \mu)$ denote the correlation neglect posterior that is generated by full correlation, that is, by repeating m times the posterior μ. In Figure 2 we plot the correlation neglect posterior generated by repeating the μ posterior m times for different m. As illustrated in the Figure, for any $\mu > 1/2$, we can find $1/2 < \mu' < \mu$ such that $\mu_{m+1}^{FC}(\mu') = \mu_m^{FC}(\mu)$. This can be done analogously for $\mu < 1/2$.

As an example, if the sender has two signals at her disposal, by using a signal that induces $\mu_0 = \sqrt{2} - 1 > 1/3$ and $\mu_1 = 2 - \sqrt{2} < 2/3$, and repeating this signal twice, the sender could generate the desired correlation neglect posteriors: $\mu^{CN}(\mu_0, \mu_0) = 1/3$ and $\mu^{CN}(\mu_1, \mu_1) = 2/3$. Such fully correlated information structure is depicted in Table 1. The entries in the table correspond to the conditional probabilities of generating those vectors of posteriors in state ω. As she is using less informative signals, the sender is able to fool the receiver with probability $\sqrt{2} - 1 > 1/3$.[8]

[8]Note that a receiver that understands that the signals are fully correlated will have correct beliefs of either μ_0 or μ_1.

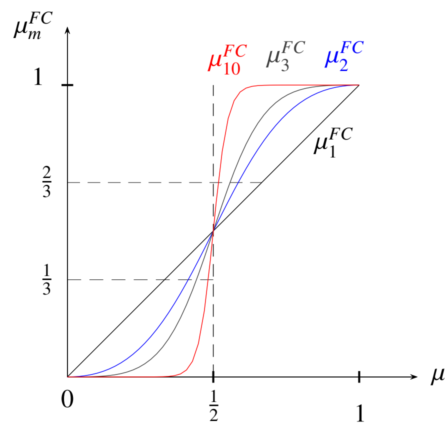


Figure 2. Amplification effect of the correlation neglect heuristic.

Table 1— Information structure with full correlation. ($m = 2, \mu_0 = \sqrt{2} - 1, \mu_1 = 2 - \sqrt{2}$)

$\tau(\cdot \mid \omega)$	(μ_0, μ_0)	(μ_1, μ_1)
$\omega = 1$	$\sqrt{2} - 1$	$2 - \sqrt{2}$
$\omega = 0$	$2 - \sqrt{2}$	$\sqrt{2} - 1$

By adding more signals and fully correlating them the sender is able to generate the desired posteriors with signals that are less informative and hence with higher probability. However, the conditional constraints impose an upper limit of $1/2$ to the probability with which the sender can fool the receiver, since any $\mu_1 > 1/2$ needs to be induced strictly more often in state $\omega = 1$ than in state $\omega = 0$.

The sender can however do better using negative correlation. Suppose for example that she has 3 signals and consider $\mu_0 = 1/3$ and $\mu_1 = 2/3$. Instead of fully correlating them, she could in state $\omega = 1$ replace the vector (μ_0, μ_0, μ_0) with the vector (μ_0, μ_0, μ_1) and its permutations. The fact that μ_0 appears more times than μ_1 in these vectors implies she will be able to generate $\mu^{CN}(\mu_0, \mu_0, \mu_1) = 1/3$. Moreover, the fact that μ_1 is also present, allows the sender to allocate less weight to (μ_1, μ_1, μ_1) while still satisfying the conditional constraints. Analogously, in state $\omega = 0$, the sender can replace the vector (μ_1, μ_1, μ_1) with the vector (μ_0, μ_1, μ_1) and its permutations. Panel A in Table 2 depicts an information structure with $m = 3$ and negative correlation. For simplicity of notation, we represent vectors of posteriors by the frequency pair (z_0, z_1), where z_i is the number of posteriors μ_i the vector contains. For example, $\boldsymbol{\mu} = (\mu_0, \mu_0, \mu_1)$ and any of its permutations will be represented by the pair $(z_0, z_1) = (2, 1)$. Using the symmetry of μ^{CN}, we denote by $\mu^{CN}(z_0, z_1)$ the correlation neglect posterior induced by any of the vectors with those frequencies. The entries in the table corresponds to the total conditional probabilities allocated to those frequency pairs. It is easy to check that the conditional constraints are satisfied, and that $\mu^{CN}(2,1) = 1/3$ and $\mu^{CN}(1,2) = 2/3$. Therefore, with just three signals, the sender can already fool the receiver half of the time![9]

Table 2— Information structures with negative correlation.

Panel A: $m = 3, \mu_0 = 1/3, \mu_1 = 2/3$

$\tau(\cdot \mid \omega)$	(3,0)	(2,1)	(1,2)	(0,3)
$\omega = 1$	0	0.5	0	0.5
$\omega = 0$	0.5	0	0.5	0

Panel B: $m = 10, \mu_0 = 0.4, \mu_1 = 0.6$

$\tau(\cdot \mid \omega)$	(10,0)	(6,4)	(4,6)	(0,10)
$\omega = 1$	0	2/3	0	1/3
$\omega = 0$	1/3	0	2/3	0

By adding more signals and negatively correlating them, the sender can generate the desired correlation neglect posteriors with a higher probability. For instance, by using $m = 10$ signals, the sender manages to fool the receiver two thirds of the time (as exemplified in Figure 1 in the introduction). She could do so by setting $\mu_0 = 0.4, \mu_1 = 0.6$ and using the information structure in Panel B of Table 2.

Continuing in this fashion, as m increases, the sender will be able to fool the receiver with a probability converging to one. The proof of Theorem 1 generalises this intuition to any number of states and any desired set of posteriors.

A. *Proof of Theorem 1*

In the proof below, we construct m homogeneous signals. Given these signals, the receiver expects to observe a multinomial distribution. Instead, the sender constructs a joint conditional distribution with support on only a subset of the realisations of the multinomial, with the aim of inducing the desired posteriors with high probability.

Specifically, each signal generates n posteriors ($n = \#\Omega$) denoted by $\{\mu^\omega\}_{\omega\in\Omega}$,[10] with the following conditional marginal distribution:

$$\begin{aligned} \tau_i^m(\mu^\omega|\omega) &= \frac{1}{n} + \frac{1}{\sqrt{m}} \equiv \alpha(m) \quad \forall\omega \in \Omega, \quad \text{i=1,...,m,} \\ \tau_i^m(\mu^\upsilon|\omega) &= \frac{1}{n} - \frac{1}{(n-1)\sqrt{m}} \equiv \beta(m) \quad \forall\omega,\upsilon \in \Omega, \quad \upsilon \neq \omega, \quad \text{i=1,...,m.} \end{aligned} \tag{4}$$

[9]Note that the information structure depicted in the table is fully revealing about the state, which implies that if the receiver understood the joint informations structure, this will result in a zero utility for the sender.

[10]To simplify the notation, when no confusion occurs, we will omit the reference to m.

The specific functional forms used for α and β are not crucial. We only require the signals to generate posteriors that span the whole space and that the signals become sufficiently uninformative as the number of signals grows so that an increasing probability can be allocated to the desired vector of posteriors. In particular any differentiable function $\alpha(m)$ such that $\alpha(m) \in (1/n, 1)$, $\beta(m) = (1 - \alpha(m))/(n-1)$, $\lim_{m\to\infty} \alpha(m) = 1/n$ and $\lim_{m\to\infty} 1/(m^2\alpha'(m)) = 0$ could be used.

Given the marginals described in (4), the posteriors $\{\mu^\omega\}_{\omega\in\Omega}$ are pinned down by Lemma 1. Intuitively, signal i generates posterior μ^ω with slightly higher probability at state ω than at any other state $\upsilon \neq \omega$. Hence, $\mu^\omega(\omega) > p(\omega)$ and $\mu^\omega(\upsilon) < p(\upsilon)$ for all $\upsilon \neq \omega$. In the limit the signals become completely uninformative as $\lim_{m\to\infty} \tau_i^m(\mu^\upsilon \mid \omega) = 1/n$ for all $\upsilon \in \Omega$.

The joint conditional distribution $\tau^m(\cdot \mid \omega)$ generates vectors of posteriors $\boldsymbol{\mu}$ with $\mu_i \in \{\mu^\upsilon\}_{\upsilon\in\Omega}$ for $i = 1, ..., m$. As in Example 1, we describe a vector of posteriors $\boldsymbol{\mu}$ in the support of $\tau^m(\cdot \mid \omega)$ by the number of times each posterior μ^υ appears in $\boldsymbol{\mu}$. Denote by $z_\upsilon(\boldsymbol{\mu})$ the proportion with which the posterior μ^υ appears in the vector $\boldsymbol{\mu}$, $z_\upsilon(\boldsymbol{\mu}) \in \{0, 1/m, ..., m/m\}$.[11]

We denote by Z_m the set of all vectors $\boldsymbol{z}(\boldsymbol{\mu}) = \{z_\upsilon(\boldsymbol{\mu})\}_{\upsilon\in\Omega}$ where $\sum_\upsilon z_\upsilon(\boldsymbol{\mu}) = 1$. Given the symmetry of the correlation neglect heuristic, if $\boldsymbol{z}(\boldsymbol{\mu}) = \boldsymbol{z}(\boldsymbol{\mu}')$ then $\mu^{CN}(\boldsymbol{\mu}) = \mu^{CN}(\boldsymbol{\mu}')$. With some abuse of notation, we denote by $\mu^{CN}(\boldsymbol{z})$ the correlation neglect posterior generated by a vector of posteriors with vector of proportions $\boldsymbol{z}$. Given Lemma 1 and Lemma 2, we have:

$$\mu^{CN}(\boldsymbol{z})(\omega) = \frac{\frac{\prod_{\upsilon\in\Omega}\mu^\upsilon(\omega)^{mz_\upsilon}}{p(\omega)^{m-1}}}{\sum_\eta \frac{\prod_{\upsilon\in\Omega}\mu^\upsilon(\eta)^{mz_\upsilon}}{p(\eta)^{m-1}}} = \frac{p(\omega)\alpha^{mz_\omega}\beta^{m(1-z_\omega)}}{\sum_\eta p(\eta)\alpha^{mz_\eta}\beta^{m(1-z_\eta)}}. \tag{5}$$

Given state $\omega \in \Omega$, we would like to generate specific vectors of proportions, $\boldsymbol{z}^\omega(m) \in Z_m$, such that $\mu^{CN}(\boldsymbol{z}^\omega(m))$ converges to ρ^ω as m increases.

Consider the case in which ρ^ω is interior, i.e. $\rho^\omega(\upsilon) > 0$ for all $\upsilon \in \Omega$.[12] To implement the posterior ρ^ω, we find the vector $\boldsymbol{\zeta}^\omega(m) = \{\zeta_\upsilon^\omega\}_{\upsilon\in\Omega}$, where $\zeta_\upsilon^\omega \in \mathbb{R}$ with $\sum_\upsilon \zeta_\upsilon^\omega = 1$, that solves:

$$\frac{p(\upsilon)\alpha^{m\zeta_\upsilon^\omega}\beta^{m(1-\zeta_\upsilon^\omega)}}{\sum_\eta p(\eta)\alpha^{m\zeta_\eta^\omega}\beta^{m(1-\zeta_\eta^\omega)}} = \rho^\omega(\upsilon) \quad \text{for all } \upsilon \in \Omega. \tag{6}$$

Rearranging we get

$$\left(\frac{\alpha}{\beta}\right)^{m\zeta_\upsilon^\omega} = \frac{\rho^\omega(\upsilon)}{p(\upsilon)}\left(\sum_\eta p(\eta)\left(\frac{\alpha}{\beta}\right)^{m\zeta_\eta^\omega}\right) \quad \text{for all } \upsilon \in \Omega, \tag{7}$$

[11]The proof can also proceed by interpreting the vectors of proportions as multinomial random variables and using the law of large numbers. We thank a referee for pointing this out.

[12]The case in which ρ^ω is not interior is addressed in the Appendix.

and by multiplying across υ and taking the n^{th} root we get

$$\left(\frac{\alpha}{\beta}\right)^{\frac{m}{n}} = \prod_{\upsilon\in\Omega}\left(\frac{\alpha}{\beta}\right)^{\frac{m\zeta_\upsilon^\omega}{n}} = \prod_{\upsilon\in\Omega}\left(\frac{\rho^\omega(\upsilon)}{p(\upsilon)}\right)^{\frac{1}{n}}\left(\sum_\eta p(\eta)\left(\frac{\alpha}{\beta}\right)^{m\zeta_\eta^\omega}\right). \tag{8}$$

By dividing (7) by (8), taking logs and rearranging we get

$$\zeta_\upsilon^\omega = \frac{1}{n} + \frac{1}{m}\frac{\ln\left(\frac{\rho^\omega(\upsilon)/p(\upsilon)}{\left[\prod_\eta \rho^\omega(\eta)/p(\eta)\right]^{\frac{1}{n}}}\right)}{\ln\left(\frac{\alpha}{\beta}\right)}. \tag{9}$$

Note that the vector $\boldsymbol{\zeta}^\omega(m)$ is not necessarily in Z_m as its entries might not be in $\{0, 1/m, ..., m/m\}$. Define by $\boldsymbol{z}^\omega(m)$ a vector in Z_m that is closest to $\boldsymbol{\zeta}^\omega(m)$ using the Euclidean distance:

$$\boldsymbol{z}^\omega(m) \in \arg\min_{z\in Z_m} |\boldsymbol{z} - \boldsymbol{\zeta}^\omega(m)|.$$

As m increases, $\lim_{m\to\infty} |\boldsymbol{z}^\omega(m) - \boldsymbol{\zeta}^\omega(m)| = 0$.[13] By the continuity of $\mu^{CN}(\cdot)$,

$$\lim_{m\to\infty} \mu^{CN}(\boldsymbol{z}^\omega(m)) = \rho^\omega.$$

Our objective is to design a joint distribution $\tau^m(\cdot \mid \omega)$ to generate the vector of proportions $\boldsymbol{z}^\omega(m)$ in state ω with high probability. However, in order to satisfy the conditional constraints (4), we also need to generate other vectors of posteriors. As we did in Example 1, we put some (small) probability on $(\mu^\upsilon, ..., \mu^\upsilon)$, $\upsilon \in \Omega$. For each $\omega \in \Omega$, consider the following joint conditional distribution:

$$\begin{aligned}
&\tau^m(\{\boldsymbol{\mu} \mid \boldsymbol{z}(\boldsymbol{\mu}) = \boldsymbol{z}^\omega(m)\} \mid \omega) = \gamma^\omega,\\
&\tau^m((\mu^\omega, ..., \mu^\omega) \mid \omega) = \lambda_\omega^\omega,\\
&\tau^m((\mu^\upsilon, ..., \mu^\upsilon) \mid \omega) = \lambda_\upsilon^\omega,\\
&\tau^m(\boldsymbol{\mu} \mid \omega) = 0 \quad \text{if } \boldsymbol{z}(\boldsymbol{\mu}) \neq \boldsymbol{z}^\omega(m), \text{ and } \boldsymbol{\mu} \neq (\mu^\upsilon, ..., \mu^\upsilon) \text{ for some } \upsilon \in \Omega,
\end{aligned} \tag{10}$$

[13]To see this, note that for sufficiently large m, we can guarantee that $0 \leq \zeta_\upsilon^\omega \leq 1$ (in fact, we show later in the proof that $\lim_{m\to\infty} \zeta_\upsilon^\omega = 1/n$). Therefore, there exists $k_\upsilon \in \mathbb{N}$, with $k_\upsilon < m$ such that $k_\upsilon/m \leq \zeta_\upsilon^\omega \leq (k_\upsilon + 1)/m$. Moreover, since $\sum_\upsilon \zeta_\upsilon^\omega = 1$, we can construct a vector of proportions $\hat{\boldsymbol{z}}$ such that $\hat{z}_\upsilon \in \{k_\upsilon/m, (k_\upsilon + 1)/m\}$. Therefore,

$$\sum_\upsilon (z_\upsilon^\omega - \zeta_\upsilon^\omega)^2 \leq \sum_\upsilon (\hat{z}_\upsilon - \zeta_\upsilon^\omega)^2 \leq \frac{n}{m^2} \quad \text{and} \quad \lim_{m\to\infty} \sum_\upsilon (z_\upsilon^\omega - \zeta_\upsilon^\omega)^2 = 0.$$

where[14]

$$\gamma^\omega = \begin{cases} \frac{\alpha}{z^\omega_\omega} & \text{if } z^\omega_\omega = 1 \\ \frac{\beta}{\max_{\upsilon\neq\omega}\{z^\omega_\upsilon\}} & \text{if } z^\omega_\omega = 0 \\ \min\left\{\frac{\alpha}{z^\omega_\omega}, \frac{\beta}{\max_{\upsilon\neq\omega}\{z^\omega_\upsilon\}}\right\} & \text{otherwise} \end{cases},$$

$$\lambda^\omega_\omega = \alpha - \gamma^\omega z^\omega_\omega,$$
$$\lambda^\omega_\upsilon = \beta - \gamma^\omega z^\omega_\upsilon.$$

In the Appendix we prove that this joint conditional distribution is well defined. Moreover, the marginals satisfy (4) for each $i = 1, ..., m$:

$$\tau^m_i(\mu^\omega|\omega) = \gamma^\omega z^\omega_\omega + \lambda^\omega_\omega = \alpha,$$
$$\tau^m_i(\mu^\upsilon|\omega) = \gamma^\omega z^\omega_\upsilon + \lambda^\omega_\upsilon = \beta, \quad \forall\upsilon \neq \omega.$$

We now show that this joint conditional distribution allocates a probability approaching one (as m increases) to the vectors of posteriors with $z(\boldsymbol{\mu}) = z^\omega(m)$. As m increases, $\alpha(m)$ and $\beta(m)$ converge to $1/n$. Using l'Hôpital's rule,

$$\begin{aligned} \lim_{m\to\infty} \frac{1/m}{\ln(\alpha(m))-\ln(\beta(m))} &= \lim_{m\to\infty} -1/m^2 \Big/ \left(\frac{1}{\alpha(m)}\frac{\partial\alpha(m)}{\partial m} - \frac{1}{\beta(m)}\frac{\partial\beta(m)}{\partial m}\right) \\ &= \lim_{m\to\infty} -1/m^2 \Big/ \left(-\frac{1}{2m^{3/2}}\frac{\alpha(m)+\beta(m)}{\alpha(m)\beta(m)}\right) \\ &= \lim_{m\to\infty} \frac{1}{m^{1/2}}\frac{1/n^2-1/m}{1/n} = 0. \end{aligned}$$

which implies that $\lim_{m\to\infty} \zeta^\omega_\upsilon = 1/n$. In addition, since $\lim_{m\to\infty} |z^\omega(m) - \boldsymbol{\zeta}^\omega(m)| = 0$, we have that $\lim_{m\to\infty} z^\omega_\upsilon = 1/n$ and $\lim_{m\to\infty} \gamma^\omega = 1$. Here we make use of the functional forms of $\alpha(m)$ and $\beta(m)$ that guarantee that the desired proportions converge to the mean and hence we can assign to the vector of such proportions a probability close to one.

Therefore, with probability converging to one, the conditional joint distribution induces a correlation neglect posterior converging to ρ^ω. This concludes the proof for an interior ρ^ω. □

B. From belief manipulation to first-best

We now formalise the sender's first-best, the lowest upper bound on her expected utility when she can freely manipulate the receiver's beliefs and freely choose from the receiver's optimal actions.

We assume that $u(a, \omega)$, $v(a, \omega)$ are continuous in a for all $\omega \in \Omega$. Given a posterior $\mu \in \Delta(\Omega)$, we denote by A_μ the set of actions in A that maximise the receiver's utility given her belief μ:

$$A_\mu = \arg\max_{a\in A} \sum_{\omega\in\Omega} \mu(\omega)u(a, \omega).$$

We assume that the correspondence $\mu \twoheadrightarrow A_\mu$ is continuous in $\Delta(\Omega)$ but for a finite set of

[14]The weight γ^ω is equally shared among all vectors of posteriors in $\{\boldsymbol{\mu} \mid z(\boldsymbol{\mu}) = z^\omega\}$

posteriors.

Define the sender's preferred actions from A_μ, $A_\mu^\omega = \arg\max_{a \in A_\mu} v(a, \omega)$. The continuity of A_μ together with the continuity of v implies (using Berge's Maximum theorem) that the function $v^\omega : \Delta(\Omega) \to \mathbb{R}$, defined as $v^\omega(\mu) = \max_{a \in A_\mu} v(a, \omega)$ is continuous in $\Delta(\Omega)$ but for a finite set of posteriors.

Let $\bar{v}^\omega = \sup_{\mu \in \Delta(\Omega)} v^\omega(\mu)$ denote the supremum of the sender's utility in state ω when she can freely choose the receiver's posterior and optimal action. We define the *first-best* for the sender as $\sum_{\omega \in \Omega} p(\omega) \bar{v}^\omega$.

Finally, we make an additional continuity assumption that allows us to show that the sender can achieve levels of utility arbitrarily close to her first-best when she has many signals. Assumption 1 ensures that even if the $\arg\sup v^\omega(\mu)$ is in the finite set of posteriors for which A_μ is not continuous, $\bar{v}^\omega$ can be approached through a sequence of posteriors beliefs:

ASSUMPTION 1: *For any $\omega \in \Omega$ and any $\mu \in \Delta(\Omega)$, there exists $a \in A_\mu^\omega$, a sequence $\{\mu_l\}_{l=1}^\infty \subset \Delta(\Omega)$ with $\mu_l \neq \mu$ for all $l = 1, ..., \infty$ and a sequence $\{a_l\}_{l=1}^\infty$ with $a_l \in A_{\mu_l}^\omega$, such that $\mu_l \to_{l \to \infty} \mu$ and $a_l \to_{l \to \infty} a$.*

Assumption 1 is weaker than lower hemicontinuity of the correspondence $\mu \twoheadrightarrow A_\mu^\omega$. To see the logic for this assumption, note that with state-independent utilities, even when the receiver's best response is not unique at some belief, one can assume that the receiver picks the sender-preferred action. In contrast, with state-dependent utility, there is no natural way to induce the receiver to choose, given some particular beliefs, different actions for different states. To guarantee that a sender's preferred action is chosen with probability close to one at each state, and even for posteriors at which $\mu \twoheadrightarrow A_\mu^\omega$ is not continuous, Assumption 1 ensures that any such action can be approached with a (potentially state-dependent) sequence of beliefs and actions. Specifically, in Example 1, $\mu \twoheadrightarrow A_\mu^\omega$ is not continuous at $\mu = 1/3$ and $\mu = 2/3$, but satisfies Assumption 1 and so we can approach $\mu = 2/3$ in state 0 with beliefs above $2/3$, and $\mu = 1/3$ in state 1 with beliefs below $1/3$.[15] All the assumptions enlisted above are satisfied if A_μ is a singleton and continuous everywhere. They are also satisfied in all the examples in this paper and the canonical examples discussed in the Bayesian Persuasion literature.

The above, together with Theorem 1, allows us to derive:

COROLLARY 1: *There exists a sequence of signal structures $\{\tau^m\}_{m \in \mathbb{N}}$ such that the sender's expected utility converges to her first-best utility.*

C. *State-independent utility and full positive correlation*

The optimal solution we construct in the proof of Theorem 1 involves negative correlation. However, in some cases, a simple correlation structure can suffice for the

[15]To see an example that violates Assumption 1 (and thus also lower hemicontinuity), consider a binary state of the world, where the receiver strictly prefers L if her belief $\mu < 1/2$, H if $\mu > 1/2$, and is indifferent among $\{L, M, R\}$ if $\mu = 1/2$. The sender's preferred action is H in state 0 as in Example 1, but M in state 1. There is however no sequence of beliefs-actions pairs (μ, a) that converges to $(1/2, M)$, as whenever $\mu < 1/2$ the receiver takes action L and whenever $\mu > 1/2$ the receivers takes action H.

sender. The next result shows that when the utility of the sender is state-independent, i.e., when she wants to induce a state-independent posterior, the sender can use full positive correlation in order to completely manipulate the receiver. Recall that we denoted by $\mu_m^{FC}(\mu) = \mu^{CN}(\mu, ..., \mu)$, the correlation neglect posterior that is generated by repeating m times the posterior μ.

PROPOSITION 1: *Given a posterior $\rho \in \Delta(\Omega)$, there exists a sequence of signal structures $\{\tau^m\}_{m\in\mathbb{N}}$ of fully positively correlated signals, such that for any $\omega \in \Omega$, and any $\epsilon > 0$,*

$$\lim_{m\to\infty} \tau^m(\{\boldsymbol{\mu} = (\mu, ..., \mu) \text{ s.t. } \tau^m(\boldsymbol{\mu}) > 0 \mid |\mu_m^{FC}(\mu) - \rho| < \epsilon\} \mid \omega) = 1.$$

Any desirable posterior ρ can be generated by using a posterior that lies in the appropriate direction away from the prior and repeating it m times. When m is large, this movement away from the prior can be minimal and can be achieved with probability close to one. As in Corollary 1, this result implies that the sender can approach her first-best (state-independent) utility.

We conclude by illustrating what the sender can achieve with full correlation and a finite number of signals, using the canonical binary utility model.

EXAMPLE 2 (The binary utility model): *Consider $\Omega = \{0, 1\}$ with prior $p = \Pr(\omega = 1) < 1/2$. The receiver takes one of two actions, $a = H$ if $\mu^{CN}(\boldsymbol{\mu}) \geq 1/2$ and $a = L$ otherwise. The sender gains 1 when the receiver chooses H and 0 otherwise, regardless of the state.*

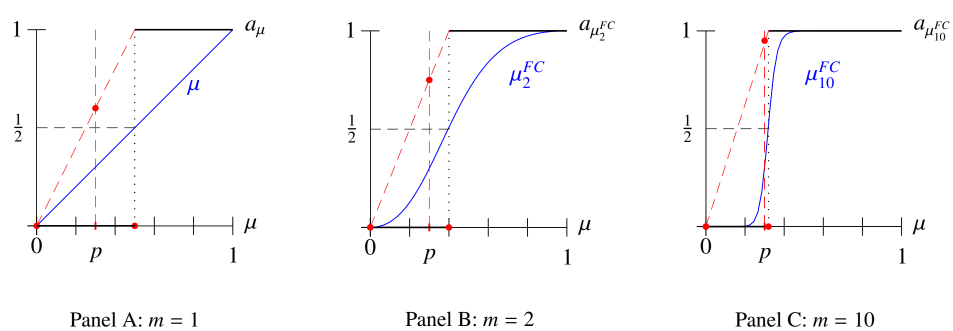


Figure 3. Optimal full correlated structure. Binary utility model.

Figure 3 illustrates the correlation neglect posterior of the receiver when the sender positively correlates all her signals. On Panel A, we illustrate the solution for one signal. The diagonal line represents the beliefs of the receiver which, for the case of one signal, corresponds to the correct beliefs. The thick horizontal lines correspond to the utility of the sender from the receiver's optimal action given the belief. Concavification of the binary utility function (the dashed line) yields a solution with two posteriors $1/2$ and 0, where posterior $1/2$ is induced with probability one in state $\omega = 1$, and with probability $p/(1-p)$ in state $\omega = 0$. This leads to expected utility of $p + (1-p)p/(1-p) = 2p$ for the sender, represented by the dot on the vertical dashed line. In Panel B and Panel C we graph $\mu_m^{FC}(\mu)$ for $m = 2$ and for $m = 10$ respectively. This corresponds to the solid S-shaped curved. Note that now the sender can induce $a = H$ whenever

$$\mu_m^{FC}(\mu) = \frac{\frac{\mu^m}{p^{m-1}}}{\frac{\mu^m}{p^{m-1}} + \frac{(1-\mu)^m}{(1-p)^{m-1}}} \geq \frac{1}{2}.$$

Concavification implies that the optimal solution would be to generate posteriors $\mu_m^{FC} = 0$ and $\mu_m^{FC} = 1/2$. To do this, the repeated signal should put weight on posteriors 0 and $\mu = p^{m-1/m}/(p^{(m-1)/m} + (1-p)^{(m-1)/m})$, inducing posterior μ with probability one in state $\omega = 1$ and probability $(p/(1-p))^{1/m}$ in state $\omega = 0$. It is easy to see that as m tends to infinity, the sender induces action $a = H$ with probability one.

III. Discussion

In this paper we show that a sender can completely manipulate a receiver with correlation neglect when she has many signals at her disposal. When the number of signals that the sender can use is bounded, she will have to use more informative signals in order to attain her desired posteriors. Using more informative signals implies however that she will face stricter constraints on the set of joint information structures she can design, due to the conditional constraints. A recent literature focuses on characterising the set of feasible joint distributions over posteriors stemming from a signal structure, sometimes in different applications (see Arieli et al. (Forthcoming), Morris (2020), Mathevet and Taneva (2019), Sandmann (2020) and Ziegler (2020)). Levy, Moreno de Barreda and Razin (2021*a*), Bergemann and Morris (2016; 2019) and Gossner (2000) suggest that having more uninformative signals relaxes the conditional constraints on information structures. Levy, Moreno de Barreda and Razin (2021*a*) also formalise how distributions over posteriors that are "closer" to full positive correlation are easier to implement. This implies that for a bounded number of signals the sender might face a trade-off between her desire to use negative correlation and relaxing the conditional constraints by using positive correlation.

Our analysis has been that of a worst-case scenario, where the receiver completely ignores correlation. In some cases, the receiver may perceive some degrees of correlation. It is beyond the scope of this paper to analyze this general case. Ellis and Piccione (2017) and Levy and Razin (2020) discuss general environments in which a decision maker perceives some degree of correlation, and one of these approaches could potentially allow an extension of our model in this direction.

REFERENCES

Alonso, Ricardo, and Odilon Câmara. 2016. "Persuading Voters." *American Economic Review*, 106(11): 3590–3605.

Arieli, Itai, Yakov Babichenko, and Fedor Sandomirskiy. 2020. "Identifiable Information Structures." *Games and Economic Behavior*, 120: 16–27.

Arieli, Itai, Yakov Babichenko, Fedor Sandomirskiy, and Omer Tamuz. Forthcoming. "Feasible Joint Posterior Beliefs." *Journal of Political Economy*.

Aumann, Robert J., and Michael B. Maschler. 1995. *Repeated Games with Incomplete Information.* Cambridge, MA: MIT Press.

Bergemann, Dirk, and Stephen Morris. 2016. "Bayes Correlated Equilibrium and The Comparison of Information Structures in Games." *Theoretical Economics*, 106: 586–591.

Bergemann, Dirk, and Stephen Morris. 2019. "Information Design: A Unified Perspective." *Journal of Economic Literature*, 57(1).

Bordley, R.F. 1982. "A Multiplicative Formula for Aggregating Probability Assessments." *Management Science*, 28(10): 1137–1148.

De Marzo, Peter, Dimitri Vayanos, and Jeffrey Zwiebel. 2003. "Persuasion Bias, Social Influence and Unidimensional Opinions." *The Quaterly Journal of Economics*, 118(3): 909–968.

Eliaz, Kfir, and Ran Spiegler. 2020. "Incentive Compatible Advertising in Non-Retail Platforms." *The RAND Journal of Economics*, 51(2): 323–345.

Eliaz, Kfir, Ran Spiegler, and Yair Weiss. Forthcoming. "Cheating with Models." *American Economic Review: Insights*.

Ellis, Andrew, and Michele Piccione. 2017. "Correlation Misperception in Choice." *American Economic Review*, 107(4): 1264–92.

Ellis, Andrew, Michele Piccione, and X Zhang. Forthcoming. "Equilibrium Securitization with Diverse Beliefs." *Theoretical Economics*.

Enke, Benjamin, and Florian Zimmermann. 2019. "Correlation Neglect in Belief Formation." *Review of Economic Studies*, 86(1): 313–332.

Eyster, Erik, and Georg Weizsäker. 2011. "Correlation Neglect in Financial Decision-Making." *Mimeo*.

Gagnon-Bartsch, Tristan, and Matthew Rabin. 2016. "Naïve Social Learning, Mislearning, and Unlearning." *Mimeo*.

Glaeser, Edward L., and Cass R. Sunstein. 2009. "Extremism and Solcial Learning." *Journal of Legal Analysis*, 1(1).

Golub, Benjamin, and Matthew O. Jackson. 2012. "How Homophily Affects the Speed of Learning and Best-Response Dynamics." *The Quarterly Journal of Economics*, 127(3): 35–71.

Gossner, Olivier. 2000. "Comparison of Information Structures." *Games and Economic Behavior*, 30(1): 44–63.

Gutmann, Sam, J. H. B. Kemperman, J. A. Reeds, and Larry A. Shepp. 1991. "Existence of Probability Measures with Given Marginals." *The Annals of Probability*, 19(4): 1781–1797.

Kallir, Ido, and Doron Sonsino. 2009. "The Neglect of Correlation in Allocation Decisions." *Southern Economic Journal*, 75(4): 1045–1066.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101(6): 2590–2615.

Levy, Gilat, and Ronny Razin. 2015*a*. "Correlation Neglect, Voting Behaviour and Information Aggregation." *American Economic Review*, 105(4): 1634–1645.

Levy, Gilat, and Ronny Razin. 2015*b*. "Does Polarization of Opinions Lead to Polarization of Platforms? The Case of Correlation Neglect." *Quarterly Journal of Political Science*, 10(3): 321–355.

Levy, Gilat, and Ronny Razin. 2020. "Combining Forecasts in the Presence of Ambiguity over Correlation Structures." *Forthcoming in the Journal of Economic Theory*.

Levy, Gilat, Inés Moreno de Barreda, and Ronny Razin. 2020. "Persuasion with Correlation Neglect." *Mimeo*.

Levy, Gilat, Inés Moreno de Barreda, and Ronny Razin. 2021*a*. "Feasible Joint Distribution of Posteriors: A Graphical Approach." *Mimeo*.

Levy, Gilat, Inés Moreno de Barreda, and Ronny Razin. 2021*b*. "Polarized extremes and the confused centre: Campaign targeting of voters with correlation neglect." *Quarterly Journal of Political Science*, 16(2): 139–155.

Mathevet, Laurent, and Ina Taneva. 2019. "Information Design." *AEJ Microeconomics*, 11(151-185).

Morris, P.A. 1977. "Combining Expert Judgments: A Bayesian Approach." *Management Science*, 23(7): 679–693.

Morris, Stephen. 2020. "No trade and Feasible Joint Posterior Beliefs." *Mimeo*.

Ortoleva, Pietro, and Erik Snowberg. 2015. "Overconfidence in Political Behavior." *American Economic Review*, 105(2): 504–35.

Sandmann, Christopher. 2020. "Recursive Information Design." *Mimeo*.

Schumacher, Heiner, and Heidi Christina Thysen. 2018. "Equilibrium Contracts and Boundedly Rational Expectation ." *Working Paper*.

Sobel, Joel. 2014. "On the Relationship Between Individual and Group Decisions." *Theoretical Economics*, 9(1): 163–185.

Ziegler, Gabriel. 2020. "Adversarial Bilateral Information Design." *Mimeo*.