Information Provision and the Curse of Knowledge

Snehal Banerjee, Jesse Davis and Naveen Gondhi*

March 27, 2023

Abstract

Better-informed individuals are typically unable to ignore their private information when forecasting others' beliefs. We study how this bias, known as "the curse of knowledge," affects costly communication and information production in a sender-receiver game. With exogenous information, cursed senders are worse communicators. However, with endogenous information production, we show that cursed senders not only produce more precise information but can, in fact, be better communicators than unbiased senders, leading to higher expected payoffs in equilibrium.

JEL Classification: D8, D9

Keywords: curse of knowledge, costly communication, information economics, expertise, diversity

*Banerjee (snehalb@ucsd.edu) is at the University of California - San Diego; Davis (Jesse_Davis@kenan-flagler.unc.edu) is at the University of North Carolina - Chapel Hill; and Gondhi (naveen.gondhi@insead.edu) is at INSEAD. We thank Judson Caskey, Archishman Chakraborty, Doron Levit, Keri Hu, Navin Kartik, Kristof Madarasz, Nadya Malenko, Yuval Rottenstreich and participants at the 12th Annual FSU SunTrust Beach Conference, the Accounting and Economics Society Webinar, the Northern Finance Association Conference and seminars at Texas A&M and University of California, Irvine, for valuable feedback. A previous version of this paper was circulated under the title "The Man(ager) Who Knew Too Much." All errors are our own.

"The single biggest problem in communication is the illusion that it has taken place." — George Bernard Shaw

"Without context, a piece of information is just a dot. It floats in your brain with a lot of other dots and doesn't mean a damn thing. Knowledge is information-in-context... connecting the dots." — Michael Ventura

1 Introduction

Coined by Camerer, Loewenstein, and Weber (1989), the curse of knowledge refers to the ubiquitous cognitive bias whereby individuals are unable to ignore their private information when forecasting the beliefs of others.[1] These mistaken beliefs are prevalent in communication because the sender, by definition, has information to share. Common intuition suggests that this bias reduces communication quality: one salient example is the struggle faced by many research professors when teaching students "simple" concepts in introductory courses, despite being thought leaders in their fields.[2] Similarly, an expert macroeconomist evaluating the impact of supply-chain constraints on potential inflation is likely to overestimate the degree to which policymakers "should have seen it coming" and so may be less inclined to voice their opinion.

In most settings, however, experts are not only responsible for sharing their expertise (e.g., communicating the relevant context) but also producing the relevant information. For instance, when designing a new seminar course, the aforementioned professor also chooses what topics to teach and which papers to include on the reading list. The macroeconomist making policy recommendations relies not only on her expertise at interpreting existing data, but also on conducting new research and analysis. Because the expert's knowledge of the context is critical for interpreting the information she produces, the natural bundling of information production and communication implies that the curse of knowledge can have important and nuanced consequences for knowledge creation and economic outcomes.

We analyze a model in which the curse of knowledge affects not only costly communication but also information production in a stylized sender-receiver game. The payoff to each player is identical and depends on the receiver's decision and the state of the world. Both players observe a public signal (henceforth, the report) of the state of the world, but in the absence of any information about the **context**, the report is imprecise.

[1] The curse of knowledge is closely related to the notion of "hindsight bias," introduced by Fischhoff (1975), which reflects the inability to correctly remember one's own priors after observing new information.

[2] In his first semester teaching at the University of Bern, Albert Einstein was able to enroll only three students in his thermodynamics course; in his second semester, his class was canceled because only one student signed up (see Grant (2018)).

The sender is privately informed about the context of this report, reflecting her expertise, and can send (a noisy version of) this information to the receiver. This costly **communication** allows the receiver to form a more precise signal from the report which improves the quality of his decision and, hence, their expected payoff. The sender can also engage in costly **information production**: she can exert effort to reduce the context-*independent* noise in the report. We assume that both communication and information production choices are covert, in that they are not directly observed by the receiver, but inferred in equilibrium.

Importantly, the sender suffers from the **curse of knowledge** and so incorrectly believes the receiver shares some of her expertise about the context of the report. We denote the degree to which the sender exhibits the curse of knowledge by the parameter ω. As ω increases, the more likely it is that the sender believes that the receiver is able to intuit her private information about the context, even in the absence of communication.

In line with common wisdom, we show that the curse of knowledge hampers communication when the level of information production is exogenously-specified. Intuitively, a cursed sender incorrectly believes that the receiver shares some of her expertise about the context and, consequently, exerting costly effort to communicate this same information is less valuable. In contrast, this same misperception leads the sender to exert more effort on information production, holding fixed the quality of communication. The sender incorrectly believes that the receiver understands the context better than he actually does and so over-estimates the weight the receiver will place on the public report. This increases the value of reducing the context-independent noise and so a cursed sender over-invests in information production.

We then characterize the conditions for existence and uniqueness of an informative equilibrium in which the sender chooses the quality of both her communication as well as the public report. We show that increases in the curse of knowledge always lead the sender to exert more effort in information production so that the context-independent noise in the report is always decreasing in ω. Paradoxically, we find that the quality of the sender's communication can *increase* in ω, depending on both the information environment and the degree to which the sender suffers from the bias.

This surprising result arises because the curse of knowledge has two offsetting effects on communication when communication quality and information production are jointly determined in equilibrium. First, as discussed above, there is a direct negative effect: an increase in ω lowers the incentive to share what the sender believes the receiver has likely intuited. The second, indirect effect is due to the complementarity between the receiver's perception of the noise in the report and the sender's choice of communication quality. Intuitively, the less noisy the report, the more attention the receiver pays and so the more value there is in providing information about the context.

We show that since information production increases with ω, this complementarity can lead to an **increase** in communication as the curse of knowledge increases. We find that when ω is sufficiently low, the indirect effect can dominate so that the quality of communication actually *increases* in the degree to which the sender suffers from the curse of knowledge. On the other hand, when ω is sufficiently high, the direct effect always dominates since the sender views communication as mostly redundant. As a result, the impact of the curse of knowledge on communication can be non-monotonic.

Our analysis implies that the impact of the curse of knowledge on the quality of the receiver's information and hence, the expected payoff, crucially depends on the expert's scope of control. In settings where the sender's primary role is to serve as a communicator, her bias leads to lower information quality and payoffs. However, the receiver may strictly prefer to employ a cursed expert, rather than an unbiased one, when information production is endogenous. In such settings, as long as the sender's bias is low, an increase in the curse of knowledge can lead to better information production *and* better communication, and so unambiguously higher payoffs for both sender and receiver.[3]

The existing empirical evidence (as we detail next) largely focuses on the impact of the curse of knowledge on communication in isolation. Our analysis provides a fresh perspective on this evidence and highlights the importance of accounting for information production in understanding the bias' impact on communication. Moreover, while our model focuses on information production, which we view as arguably most natural, the economic mechanism underlying our results would apply to other essential activities which complement expert communication. Given the widespread evidence of this bias in experts, our analysis yields insights on how the curse of knowledge can have unexpected, but far-reaching, impact on real outcomes.

1.1 Empirical relevance and related literature

The curse of knowledge is an aspect of "perspective taking" that has been widely studied by psychologists and anthropologists.[4] The bias has been widely documented and arises at any age, across different cultures, and in a variety of settings and information environments (see the surveys by Hawkins and Hastie (1990), Blank, Musch, and Pohl (2007) and Ghrear, Birch, and Bernstein (2016), and the papers detailed within). There is also ample

[3]More generally, as long as information production is sufficiently sensitive to the sender's bias, the expected payoff can be increasing in ω even as communication quality declines.

[4]As highlighted by Nickerson (1999), an individual engaging in perspective taking (or "putting themselves in someone else's shoes") finds it difficult to imagine that others do not know what he knows. This is what gives rise to the "curse of knowledge."

evidence that a range of communication methods can give rise to the curse of knowledge: while the original research focused on written communication (e.g., Fischhoff, 1975), there is substantial evidence that individuals exhibit the curse of knowledge with respect to oral communication (Keysar, 1994), graphical messages (Xiong, van Weelden, and Franconeri, 2019) and visual illustrations (Bernstein, Atance, Loftus, and Meltzoff, 2004).

Most importantly, the literature documents that experts are particularly susceptible to the curse of knowledge. For instance, Arkes, Wortmann, Saville, and Harkness (1981) show that physicians given both symptoms and the "correct" diagnosis, overestimate the likelihood that a physician presented with the symptoms (only) would correctly diagnose the ailment. In Anderson, Jennings, Lowe, and Reckers (1997), judges who evaluate the quality of an auditors' ex-ante decision are influenced by their ex-post knowledge of the outcome. Kennedy (1995) shows that both auditors and MBA students are subject to the curse of knowledge when they evaluate ex-ante forecasts using ex-post bankruptcy outcomes. Finally, there is substantial evidence that traditional methods of debiasing have limited, if any impact: a series of papers (see Pohl and Hell (1996), Kennedy (1995), and the survey by Harley (2007)) show that even individuals with prior experience, who receive feedback on their performance and are accountable for their actions, and who are provided with direct warnings about the bias still exhibit the curse of knowledge.

Camerer et al. (1989) are the first to explore the implications of the curse of knowledge in economic decision-making. Using an experimental design, they find that this bias is a robust feature of individual forecasts and is not eliminated by incentives or feedback. The authors conclude that the curse of knowledge can help "alleviate the inefficiencies that result from information asymmetries" because better-informed agents do not fully exploit their information advantage in a competitive setting. Our analysis generates new insight in part due to the cooperative nature of the underlying game. In our setting, a biased sender also fails to capitalize on her information advantage, but this implies that the signals she sends to the receiver and, hence, the expected payoff can be less efficient.

Our paper is most closely related to Madarász (2011).[5] He shows that when the receiver exhibits the curse of knowledge (or "information projection") and evaluates the sender's expertise, the sender overproduces information that is a substitute for the receiver's ex-post information and under-produces complementary information. When the sender is biased and communication is costly, Madarász (2011) shows that the sender speaks too rarely and is difficult to understand. Our analysis complements this work. With exogenous information production, we also show that biased experts poorly communicate their private information;

[5] In another related paper, Madarász (2015) shows that in a persuasion game with costly verification and a biased receiver, the equilibrium may feature credulity or disbelief.

However, a biased sender chooses to *reduce* the noise in a public report. Moreover, when the sender engages in both communication and information production, the complementarity *across* these activities implies that the curse of knowledge can improve the quality of communication, which is distinct from earlier work. Similarly, as in Madarász (2011), we show that increasing the sender's expertise amplifies the effect of the curse of knowledge. To our knowledge, however, we are the first to consider how the receiver's expertise and the diversity of their information affects the impact of the sender's biased beliefs.

2 Model Setup

There are two players: a sender S (she) and a receiver R (he). Their preferences are perfectly aligned and the payoff is given by

$$U = -(\theta - k)^2, \tag{1}$$

where θ reflects the state of the world, and k denotes the receiver's action. We assume that θ is normally distributed with mean zero and precision τ, i.e.,

$$\theta \sim N(0, 1/\tau). \tag{2}$$

The sender can pay an effort cost $c(p)$ to produce a commonly-observed signal

$$s = \theta + x + \varepsilon, \tag{3}$$

where $\varepsilon \sim N(0, 1/p)$ and $x \sim N(0, 1/\tau_x)$. The additional source of noise, x, captures the role of **context** for interpreting the signal. Having more precise information about the context, x, makes the signal, s, more informative about the fundamental state θ. We assume that x is privately known to the sender, reflecting her expertise and understanding of the underlying environment. The sender can share her expertise with the receiver by paying an effort cost $\kappa(q)$ to generate a commonly-observed signal

$$y = x + \eta, \tag{4}$$

where $\eta \sim N(0, 1/q)$ is independent of all other random variables.[6]

The choice of $p > 0$ reflects the sender's **information production** decision; the choice

[6]We focus on normally-distributed random variables and a quadratic loss function for analytic tractability and ease of exposition. This also allows for a transparent comparison of our results to existing analyses of sender-receiver games.

of $q > 0$ reflects the **quality** of the sender's **communication**. We assume that the sender's production and communication are covert: the receiver does not observe the choice of p and q but forms a conjecture, $\hat{p}$ and $\hat{q}$, that the sender takes as given when making her choices. The cost of information production $c(\cdot)$ and communication $\kappa(\cdot)$ are increasing and convex. We will say that the cost function is "well-behaved" if $c(0) = c'(0) = 0$ and $\kappa(0) = \kappa'(0) = 0$.

The random variables θ, ε, x, and η are independent of each other. Both the sender and receiver have common priors about the joint distribution of these random variables which are consistent with the objective joint distribution. Let $\mathbb{E}_i[\cdot]$ and $\mathbb{V}_i[\cdot]$ denote the conditional expectation and conditional variance of player $i \in \{S, R\}$, given their information sets, i.e.,

$$\mathbb{E}_R[\theta] = \mathbb{E}[\theta|s, y], \quad \mathbb{V}_R[\theta] = \mathbb{V}[\theta|s, y], \text{ and} \tag{5}$$

$$\mathbb{E}_S[\theta] = \mathbb{E}[\theta|s, y, x], \quad \mathbb{V}_S[\theta] = \mathbb{V}[\theta|s, y, x]. \tag{6}$$

Importantly, the sender suffers from the **curse of knowledge**: she mistakenly believes that the receiver shares some of her expertise. In particular, consider a "truth or nothing" signal about x,

$$z = \begin{cases} x & \text{with probability } \omega \\ \emptyset & \text{with probability } 1-\omega \end{cases}, \tag{7}$$

where the signal z reveals x perfectly with probability ω and nothing otherwise. We model the curse of knowledge by assuming that the seller believes that the receiver also observes the signal z when, in reality, he does not. The parameter $\omega \in [0, 1)$ measures the degree to which the sender exhibits the curse of knowledge. If $\omega = 0$, the sender exhibits rational expectations; otherwise, as ω increases, the sender's forecast about the receiver's beliefs is biased toward her own. In particular, note that

$$\mathbb{E}_S[\mathbb{E}_R[\theta]] = \mathbb{E}[\mathbb{E}[\theta|s, y, z]|\, s, y, x] = (1-\omega)\mathbb{E}_R[\theta] + \omega\mathbb{E}_S[\theta]. \tag{8}$$

Note that the characterization of conditional expectations in (8) corresponds to the formulation first utilized by Camerer, Loewenstein, and Weber (1989). Our specification is also consistent with the notion of "information projection" developed by Madarász (2011).

The timing of the game is as follows:

- Date $t = 0$: The sender optimally exerts effort p for information production and effort q for communication to maximize her expected payoff i.e.,

$$p, q = \arg\max_{p,q} \mathbb{E}_S\left[-\left(\theta - \hat{k}_R\right)^2\right] - c(p) - \kappa(q), \tag{9}$$

where $\hat{k}_R$ reflects the sender's forecast of the receiver's action.

- Date $t = 1$: Nature draws the random variables $\{\theta, \varepsilon, \eta, x, z\}$. The signals s and y are commonly observed; x is privately observed by the sender. The receiver chooses action k_R to maximize her expected payoff, i.e.,

$$\max_k \mathbb{E}_R \left[-(\theta - k)^2 \right], \tag{10}$$

given $\hat{p}$ and $\hat{q}$, which denote the receiver's inference of the sender's information production and communication decisions.

- Date $t = 2$: Players receive the payoff $-(\theta - k_R)^2$.

We assume that the sender is naive about her behavioral bias in that her choices do not "correct" for the fact that she is subject to the curse of knowledge. This is consistent with empirical evidence which implies that individuals continue to exhibit the curse of knowledge even if they are made aware of the fact that they are doing so (see Section 1.1.)[7]

3 Equilibrium

In our model, there always exists an uninformative equilibrium in which the sender chooses to exert no effort in information production or communication (i.e., chooses $p = 0$, $q = 0$), and a no-communication equilibrium in which the sender chooses to exert no effort in communication, but some effort in information production (i.e., chooses $q = 0$ and $p > 0$).[8] The former equilibrium is trivial and the latter equilibrium is isomorphic to the one-dimensional analysis in Section 3.2. As such, in what follows, we focus on the informative equilibrium in which the sender chooses to both produce an informative report and communicate her private information (i.e., $p, q > 0$) and the receiver's optimal action depends on s and y.

[7]It would be interesting to study how a sophisticated sender could commit to an information provision strategy which accounted for, and possibly exploited, her curse of knowledge but we leave this for future work.

[8]The former is an equilibrium because, the receiver's conjecture ($\hat{p} = \hat{q} = 0$) implies that he places no weight on the signals (s and y, respectively) and so the sender has no incentive to reduce the noise. The latter is an equilibrium because, even absent any communication ($q = 0$), if the receiver conjectures $\hat{p} > 0$, he places positive weight on the sender's report and so the sender exerts effort.

3.1 Preliminary Analysis

We solve the model by working backwards. Given the receiver's objective function in (10), his optimal action is given by

$$k_R^* = \mathbb{E}_R[\theta] = \mathbb{E}[\theta|s,y]\,. \tag{11}$$

Using standard properties of normal distributions, we can express this as

$$k_R^* = \Lambda\left(s - \mu_x(y)\right), \text{ where } \mu_x(y) \equiv \mathbb{E}[x|y] = \lambda y \tag{12}$$

and where

$$\lambda \equiv \frac{\hat{q}}{\tau_x + \hat{q}}, \quad \Lambda \equiv \frac{\tau_s}{\tau + \tau_s}, \quad \text{and} \quad \frac{1}{\tau_s} \equiv \frac{1}{\hat{p}} + \frac{1}{\tau_x + \hat{q}}. \tag{13}$$

and where $\hat{p}$ and $\hat{q}$ reflect the receiver's inference about the sender's choices.

However, given the sender's curse of knowledge, her inference about the receiver's action is given by

$$\hat{k}_R = \mathbb{E}[\theta|s,y,z] = \begin{cases} \Lambda\left(s - \mu_x(y)\right) & \text{with prob } 1-\omega \\ \Lambda_\omega(s-x) & \text{with prob } \omega \end{cases}, \tag{14}$$

where

$$\Lambda_\omega = \frac{\tau_\omega}{\tau + \tau_\omega}, \quad \text{and} \quad \frac{1}{\tau_\omega} = \frac{1}{\hat{p}}. \tag{15}$$

The difference in the perceived precisions, $\tau_\omega > \tau_s$, captures the impact of the curse of knowledge: the sender mistakenly believes that the receiver has a better understanding of the context and so over-estimates his understanding of the report, s.

Given her beliefs about the receiver's action, the sender optimally chooses the quality of communication q and the precision of information produced p to maximize the objective

$$\begin{aligned} u_S(p,q,\hat{p},\hat{q}) &\equiv \mathbb{E}_S\left[-\left(\theta - \hat{k}_R\right)^2\right] \\ &= -\mathbb{E}\left[\ (1-\omega)\left(\theta - \Lambda\left(s-\mu_x(y)\right)\right)^2 \ + \omega\left(\theta - \Lambda_\omega(s-x)\right)^2\right] \end{aligned}$$

subject to the cost $c(p) + \kappa(q)$. The sender takes the receiver's conjectured $\hat{p}$ and $\hat{q}$ as given, and so the marginal benefit of increasing her effort along either dimension can be written as:

$$\frac{\partial}{\partial p} u_S(p,q,\hat{p},\hat{q}) = \frac{1}{p^2}\left(\omega\Lambda_\omega^2 + (1-\omega)\Lambda^2\right), \tag{16}$$

and

$$\frac{\partial}{\partial q} u_S \left(p, q, \hat{p}, \hat{q}\right) = \left(1 - \omega\right) \Lambda^2 \frac{\lambda^2}{q^2}. \tag{17}$$

In a perfect Bayesian equilibrium, the receiver's conjectured choices, $\hat{p}$ and $\hat{q}$ will be the same as the sender's choice. The receiver's expected payoff, given p and q, is the same as the expected payoff under the objective distribution:

$$u_R \left(p, q\right) \equiv \mathbb{E}\left[-\left(\theta - k_R^*\right)^2 \big|_{\hat{p}=p, \hat{q}=q} \right] = -\, \mathbb{V}\left(\theta | s, y\right)\big|_{\hat{p}=p, \hat{q}=q} \tag{18}$$

which is increasing in the precision of both signals since,

$$\frac{\partial}{\partial p} u_R \left(p, q\right) = \frac{\Lambda^2}{p^2} > 0 \quad \text{and} \quad \frac{\partial}{\partial q} u_R \left(p, q\right) = \frac{\Lambda^2}{\left(\tau_x + q\right)^2} > 0. \tag{19}$$

3.2 Equilibrium: One Dimension

We first study how the curse of knowledge affects the equilibrium precision of one signal when the noise in the other signal is exogenously specified. Underlying these results is the sender's mistaken perception of the receiver's information about the context, x: suffering from the curse of knowledge, she believes that the receiver may be able to infer her private information, even in the absence of any communication.

Proposition 1. *If information production is fixed (i.e., $p > 0$ is exogenous) and the cost-function is well-behaved, then there is a unique informative (i.e., $q^* > 0$) equilibrium in which endogenous communication quality (q^*) and, hence, the expected payoff, is decreasing in the curse of knowledge (ω).*

Intuitively, if the receiver is able to partially infer x, then the sender's communication of this same information is less valuable. The more biased the sender with respect to the receiver's beliefs (i.e., the larger ω), the less value she perceives from exerting effort (see equation (17)): in equilibrium, communication quality decreases with the curse of knowledge. Moreover, the decrease in communication quality reduces the precision of the receiver's forecast, which reduces the expected payoff. This result is in line with the narrative from the psychology literature which suggests that cursed experts tend to communicate poorly and do not exert much effort in "making the case clearly" because they over-estimate the extent to which their audience is "on the same page."

In contrast, the next result shows that the curse of knowledge can improve information production.

Proposition 2. *If communication quality is fixed (i.e., $q > 0$ is exogenous) and the cost-function is well-behaved, then there is a unique informative (i.e., $p^* > 0$) equilibrium in which endogenous information production (p^*) and, hence, the expected payoff, is increasing in the curse of knowledge (ω).*

Suffering from the curse of knowledge, the sender mistakenly believes that the receiver has a more precise understanding of the context. Knowing such information would allow the receiver to place a higher weight ($\Lambda_\omega > \Lambda$) on the commonly-observed report, s, which increases the benefit of producing a more precise report (see equation (16)). This effect is increasing in the extent of the sender's bias, and in equilibrium, p^* is increasing in ω, in contrast to the setting with endogenous communication.

In summary, the impact of the curse of knowledge depends crucially on the nature of the information the sender is choosing to convey. Though communication is a substitute for the private information the sender mistakenly believes the receiver already knows, the report which she produces is a complement. As a result, while the sender has weaker incentives to communicate her private information about the context, she has a stronger incentive to produce a more informative report. The next section characterizes how these two channels interact when the sender endogenously chooses both the quality of communication and the precision of information production.

3.3 Equilibrium: Two Dimensions

We now analyze how the curse of knowledge affects both information production and communication jointly.

Lemma 1. *The marginal benefit of higher quality communication is increasing in the receiver's conjectured precision of the sender's report ($\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial q \partial \hat{p}} > 0$); the marginal benefit of increased information production is increasing in the conjectured quality of the sender's communication ($\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial p \partial \hat{q}} > 0$).*

In equilibrium, when the receiver is more informed about the context, he places more weight on the sender's report. Thus, when $\hat{q}$ is high, the higher the sender's value of increasing the report's precision. Similarly, if the report is more informative in equilibrium (i.e., if $\hat{p}$ is high), the sender believes it is more valuable to communicate more precisely. In equilibrium, this complementarity can be transformational: the relationship between the curse of knowledge and communication can be positive, counter to the standard intuition.

While the sender's optimal p^* and q^* are unique given the receiver's conjectured $\hat{p}$ and $\hat{q}$, the complementarity established in Lemma 1 also implies that there may be multiple equilibria. In the following proposition, we establish that, as long as ex-ante, fundamental uncertainty is sufficiently low, the complementarity between the optimal p and q is relatively small and so a unique informative equilibrium exists.

Proposition 3. *If both information production and communication quality are chosen by the sender, then there exists a unique informative equilibrium ($p^*, q^* > 0$) if the cost function is well-behaved and $\tau > 1$. In this equilibrium, (i) the precision of the report, p^*, is always increasing in the curse of knowledge; (ii) there exist threshold $\bar{\omega} \in (0,1)$ such that if $\omega > \bar{\omega}$, the quality of communication, q^*, is decreasing in the curse of knowledge.*

When τ is sufficiently high, the equilibrium weight the receiver places on each signal is less sensitive to an increase in the precision chosen by the sender. This rules out multiple equilibria, e.g., when communication is higher, the marginal benefit of increasing information production is outweighed by the decline in the marginal benefit of communication itself.

Though the value of information production (p) is lower when communication is lower (q), in equilibrium, the positive direct effect of the curse of knowledge on information production *always* outweighs the indirect effect of any reduction in communication. As a result, the more the sender suffers from the curse of knowledge, the more precise her report.

The effect on communication, however, is more nuanced. The key factor is the sender's perception of the weight the receiver places on the report: the impact of the curse of knowledge on the marginal benefit of information production is increasing in the difference between Λ_ω and Λ.[9] When this difference is sufficiently large, p increases more quickly which has the indirect effect of making *communication* more valuable, potentially flipping the sign of $\frac{dq}{d\omega}$: communication can actually increase in the sender's bias. This can only occur, however, when the curse of knowledge is not too large (ω must be sufficiently small): if not, the sender believes that communication has little value (since the receiver is more likely to know the context), even if the report is very precise.

To illustrate these effects, we consider a setting in which the cost of information production is given by $c(p) = c_p p$ while the cost of improving the quality of communication is given by $\kappa(q) = \kappa_q q$, where $c_p, \kappa_q > 0$. While such a cost-function is not "well-behaved" (since both $c'(0), \kappa'(0) > 0$), a unique equilibrium with both communication and information production exists as long as κ_q and c_p are sufficiently small. Moreover, this specification is analytically tractable, allowing us to characterize p^* and q^* in closed-form.

[9]The difference between Λ_ω and Λ is largest when (i) the context is sufficiently important (τ_x is sufficiently low) and (ii) ex-ante uncertainty (τ) is neither too low or too high. If τ is too high, then both Λ_ω and Λ tend to zero; if τ is too low, then both Λ_ω and Λ tend to one.

Lemma 2. *(Linear Cost Equilibrium) If $c_p < \bar{c}$ and $\kappa_q < \bar{\kappa}$, derived in the proof, then in the unique informative equilibrium, p^* and q^* are given by*

$$p^* = G - \tau, \tag{20}$$

$$q^* = \left(\frac{1}{\psi} - \tau\right)\left(1 - \frac{\tau}{G}\right) - \tau_x \tag{21}$$

where $\psi = \sqrt{\frac{\kappa_q}{1-\omega}}$ and $G = \sqrt{\frac{1+\kappa_q\tau\left(\tau - \frac{2}{\psi}\right)}{c_p}}$.
(i) The precision of the report, p^*, is always increasing in the curse of knowledge;
(ii) The quality of communication, q^*, is increasing in the curse of knowledge if

$$\frac{\kappa_q\tau^2}{c_p\psi G^3}(1 - \tau\psi) > 1 - \frac{\tau}{G} \tag{22}$$

(iii) The expected payoff is increasing in the curse of knowledge if

$$\frac{\kappa_q\tau}{c_p\psi^2 G^3}(1 - \tau\psi) > 1 - \frac{\tau}{G} \tag{23}$$

Figure 1 illustrates how the curse of knowledge affects the equilibrium precision choices and, consequently, alters the expected payoff. Note that, in panel (a), the precision of the report (p) is always increasing in curse of knowledge. Moreover, because information production increases sufficiently with ω in this numerical example, so does optimal communication quality (q) when the sender is not too biased. However, there exists a threshold $\bar{\omega} \approx 0.5$ such that if $\omega > \bar{\omega}$, the quality of communication, q^*, is decreasing in ω.

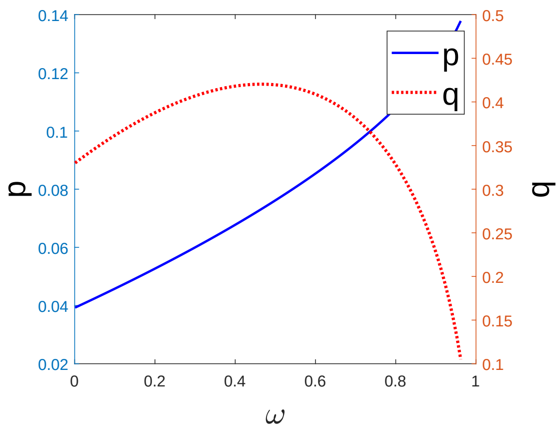
As panel (b) illustrates, this leads to non-monotonicity in the expected payoff as the curse of knowledge changes. For small ω, an increase in ω increases both p and q, which both increases the expected payoff. This is true even as q begins to fall (since p continues to increase) while for sufficiently high ω (greater than ≈ 0.85), the decline in communication eventually leads to a lower expected payoff.

4 Discussion and Concluding Remarks

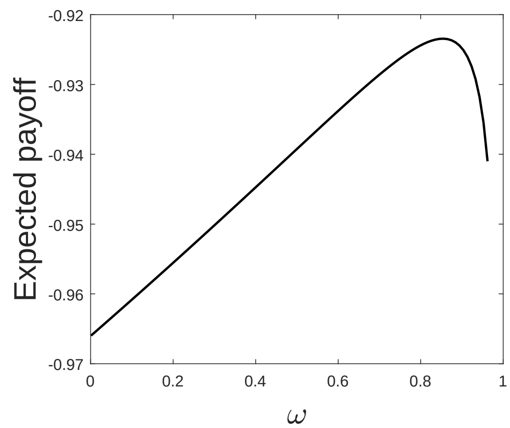
In Online Appendix B.1, we extend the model so that both the sender and receiver have (partial) understanding of the context and analyze how the impact of the curse of knowledge is affected by receiver expertise and diversity.[10] In equilibrium, we show both communication quality and information production improve with receiver expertise and diversity: when the

[10]Diversity measures the dispersion in the sender and receiver's beliefs, holding fixed their expertise.

Figure 1: Optimal communication and information production with linear costs
The figure plots the choice of message precision q (dashed), information production p (solid), and the expected payoff $\mathbb{E}[V(R,k)]$ under costly communication, as a function of the curse of knowledge. The other parameters of the model are $\tau = 1, \tau_x = 0.01$, $\kappa_q = 0.01$ and $c_p = 0.75$.



(a) q (dashed) and p (solid) vs ω



(b) Expected value vs ω

receiver has more expertise, he is better able to interpret the information in the report while an increase in diversity implies that, all else equal, the sender's information is incrementally more informative for the receiver.

On the other hand, higher receiver expertise and more diversity also increases the value of the information the sender *incorrectly* believes the receiver infers. As a result, increases in expertise and diversity amplify the impact of the sender's curse of knowledge on her perception of the value of communication and information production. For example, when information production is fixed, this implies that communication falls more quickly with diversity as the curse of knowledge grows, consistent with common intuition. However, we show that the level of communication chosen by a diverse sender remains higher, highlighting the importance of studying the equilibrium effects.

Our specification of the receiver's objective highlights the crucial role that expertise plays in converting data (in our setting, the report) into actionable information. In particular, the more precise the sender's communication, the more valuable the report. In Online Appendix B.2, we consider a generalized setting in which the objective is $\beta\theta + (1-\beta)x$, i.e., the sender's communication does not simply complement the report but is also a (partial) substitute. For β sufficiently high, the results in our main model continue to hold; however, when β is small, the curse of knowledge always hampers both communication and information production. These results underscore the role of complementarity in generating our key insight.

Finally, we assume both that (i) the sender's effort choices for communication and information production are covert (i.e., unobservable to the receiver) and (ii) that the sender commits to the quality of communication (q) without conditioning on the realization of the context (x). However, in Online Appendix B.3, we show that the equilibrium precision choices are unchanged when we relax either assumption.[11]

In summary, we study how the curse of knowledge affects information provision (i.e., not only costly communication but also information production) and expected payoffs in a sender-receiver game. While the sender's bias reduces the perceived value of communication, it increases her incentive to produce a more precise report. When both communication and production are endogenous, this latter channel can reverse the former effect: counter to standard intuition, the quality of communication can increase when the sender exhibits the curse of knowledge due to the complementary role played by these two sources of information. Our results suggest that the impact of the curse of knowledge on information provision and other actions is nuanced and crucially depends on how these decisions interact with the expert's incentives to communicate.

[11] Our assumption that the choice of q is not observable also precludes the possibility that the sender could use her choice of q to signal her knowledge of x if q were chosen after x is observed.

References

John C. Anderson, Marianne M. Jennings, D. Jordan Lowe, and Philip M.J. Reckers. The mitigation of hindsight bias in judges' evaluation of auditor decisions. *Auditing*, 16(2): 20–39, 1997.

Hal R. Arkes, Robert L. Wortmann, Paul D. Saville, and Allan R. Harkness. Hindsight bias among physicians weighing the likelihood of diagnoses. *Journal of Applied Psychology*, 66 (2):252–254, 1981.

Daniel M. Bernstein, Cristina Atance, Geoffrey R. Loftus, and Andrew Meltzoff. We saw it all along: Visual hindsight bias in children and adults. *Psychological Science*, 15(4): 264–267, 2004.

Hartmut Blank, Jochen Musch, and Rüdiger F. Pohl. Hindsight bias: On being wise after the event. *Social Cognition*, 25(1):1–9, 2007.

Colin Camerer, George Loewenstein, and Martin Weber. The curse of knowledge in economic settings: An experimental analysis. *Journal of Political Economy*, 97(5):1232–1254, 1989.

Baruch Fischhoff. Hindsight $\neq$ foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1 (3):288–299, 1975.

Siba E Ghrear, Susan AJ Birch, and Daniel M Bernstein. Outcome knowledge and false belief. *Frontiers in psychology*, 7:118, 2016.

Adam Grant. Those who can do, can't teach. *The New York Times*, https://nyti.ms/2P5QedB, August 25, 2018 2018.

Erin M. Harley. Hindsight bias in legal decision making. *Social Cognition*, 25(1):48–63, 2007.

Scott A Hawkins and Reid Hastie. Hindsight: Biased judgments of past events after the outcomes are known. *Psychological bulletin*, 107(3):311, 1990.

Jane Kennedy. Debiasing the curse of knowledge in audit judgment. *The Accounting Review*, 70(2):249–273, 1995.

Boaz Keysar. The illusory transparency of intention: Linguistic perspective taking in text. *Cognitive Psychology*, 26(2):165–208, 1994.

Kristóf Madarász. Information projection: Model and applications. *The Review of Economic Studies*, 79(3):961–985, 2011.

Kristóf Madarász. Projection equilibrium: Definition and applications to social investment and persuasion. 2015.

Raymond S Nickerson. How we know—and sometimes misjudge—what others know: Imputing one's own knowledge to others. *Psychological bulletin*, 125(6):737, 1999.

Rüdiger F. Pohl and Wolfgang Hell. No reduction in hindsight bias after complete information and repeated testing. *Organizational Behavior and Human Decision Processes*, 67(1):49–58, 1996.

Cindy Xiong, Lisanne van Weelden, and Steven Franconeri. The illusory transparency of intention: Linguistic perspective taking in text. *IEEE Transactions on Visualization and Computer Graphics*, 2019. doi: 10.1109/TVCG.2019.2917689.

A Proofs

Proof of Proposition 1.

Proof. (1) An equilibrium without informative communication ($\hat{q} = q^* (\hat{q}) = 0$) always exists. If the receiver conjectures that $\hat{q} = 0$, the sender's best-response to this conjecture is $q^* (\hat{q} = 0) = 0$, since (i) the receiver does not condition on the signal y when $\hat{q} = 0$ and (ii) effort (choosing $q > 0$) is always costly. This confirms the receiver's conjecture and constitutes an equilibrium.

(2) A unique informative equilibrium ($\hat{q} = q^* (\hat{q}) > 0$) always exists. We prove this in two steps. First, it is always the case that, given the receiver's conjecture, $\hat{q}$, the seller has a unique best-response, $q^* (\hat{q})$. This is because (i) $\kappa (q)$ is increasing and convex and (ii) $\frac{\partial u_S(p,q,\hat{p},\hat{q})}{\partial q} > 0$ while $\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial q^2} = \left(-\frac{2}{q}\right) \frac{\partial u_S(p,q,\hat{p},\hat{q})}{\partial q} < 0$. If the cost function is well-behaved, then $\kappa' (0) = \kappa (0) = 0$, which implies that $q^* (\hat{q}) > 0$ is the unique solution which solves $\frac{\partial u_S(p,q,\hat{p},\hat{q})}{\partial q}|_{q=\hat{q},p=\hat{p}} = \kappa' (q)$.

Second, in equilibrium, the receiver's conjecture is correct, i.e., $\hat{q} = q^* (\hat{q})$. This constitutes the unique informative equilibrium as long as there is a single $\hat{q} = q^* (\hat{q}) = q > 0$ which solves the sender's first-order condition. Since $\kappa (q)$ is increasing, this implies that is sufficient to show that

$$\frac{\partial}{\partial q} \left(\frac{\partial u_S (p = \hat{p}, q, \hat{p}, \hat{q})}{\partial q} \right) = \frac{\partial}{\partial q} \left(\left. \frac{\lambda^2 \Lambda^2 (1 - \omega)}{q^2} \right|_{\hat{q}=q} \right) \tag{24}$$

$$= -\frac{2p^2 (1 - \omega) (p + \tau)}{(p (q + \tau + \tau_x) + \tau (q + \tau_x))^3} < 0, \tag{25}$$

which always holds.

(3) The equilibrium communication quality $q^* (\hat{q}) > 0$ is decreasing in ω since

$$\frac{\partial}{\partial \omega} \left(\frac{\partial u_S (p = \hat{p}, q, \hat{p}, \hat{q})}{\partial q} \right) = \frac{\partial}{\partial \omega} \left(\left. \frac{\lambda^2 \Lambda^2 (1 - \omega)}{q^2} \right|_{\hat{q}=q} \right) < 0 \tag{26}$$

□

Proof of Proposition 2.

Proof. (1) An uninformative equilibrium in which ($\hat{p} = p^* (\hat{p}) = 0$) always exists. The proof is analogous to the proof found in Proposition 1 .

(2) A unique informative equilibrium ($\hat{p} = p^* (\hat{p}) > 0$) always exists. The proof is analogous to the proof found in Proposition 1. First, since $\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial p^2} = -\frac{2}{p} \frac{\partial u_S(p,q,\hat{p},\hat{q})}{\partial p} < 0$,

the optimal $p^*(\hat{p})$ is the solution to $\frac{\partial u_S(p,q,\hat{p},\hat{q})}{\partial p} = c'(p)$. Second, it is sufficient to ensure that there is a unique informative equilibrium in which $\hat{p} = p^*(\hat{p}) = p > 0$ as long as

$$\frac{\partial}{\partial p}\frac{\partial u_S(p = \hat{p}, q, \hat{p}, \hat{q})}{\partial p} = \frac{\partial}{\partial p}\left(\left.\frac{(1-\omega)\Lambda^2 + \omega\Lambda_\omega^2}{p^2}\right|_{\hat{p}=p}\right) \tag{27}$$

$$= -\frac{2(1-\omega)(q+\tau_x)^2(q+\tau+\tau_x)}{(p(q+\tau+\tau_x)+\tau(q+\tau_x))^3} - \frac{2\omega}{(p+\tau)^3} < 0, \tag{28}$$

which always holds.

(3) The equilibrium information production $p^*(\hat{p}) > 0$ is increasing in ω since

$$\frac{\partial}{\partial \omega}\left(\frac{\partial u_S(p = \hat{p}, q, \hat{p}, \hat{q})}{\partial p}\right) = \frac{\partial}{\partial \omega}\left(\left.\frac{(1-\omega)\Lambda^2 + \omega\Lambda_\omega^2}{p^2}\right|_{\hat{p}=p}\right) \tag{29}$$

$$= \frac{p\tau(p(2q+\tau+2\tau_x)+2\tau(q+\tau_x))}{(p+\tau)^2(p(q+\tau+\tau_x)+\tau(q+\tau_x))^2} > 0. \tag{30}$$

$\square$

Proof of Lemma 1

Proof. This follows immediately from differentiation of equations (16) and (17):

$$\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial p \partial \hat{q}} = \frac{2(1-\omega)\tau}{p^2}\Lambda^3\left(\frac{1}{\tau_x+\hat{q}}\right)^2 > 0, \tag{31}$$

$$\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial q \partial \hat{p}} = \frac{2(1-\omega)\Lambda^3}{\hat{p}^2}\left(\frac{\hat{q}}{\hat{q}+\tau_x}\right)^2\frac{1}{q^2} > 0. \tag{32}$$

Of course, in equilibrium ($p = \hat{p}, q = \hat{q}$), these equations exactly correspond, since

$$\left.\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial p \partial \hat{q}}\right|_{\hat{p}=p,\hat{q}=q} = \left.\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial \hat{p} \partial q}\right|_{\hat{p}=p,\hat{q}=q} = \frac{2p\tau(1-\omega)(q+\tau_x)}{(p(q+\tau+\tau_x)+\tau(q+\tau_x))^3}. \tag{33}$$

$\square$

Proof of Proposition 3

Proof. **Part 1 (Equilibrium Existence and Uniqueness)** First, we establish that, given the receiver's conjecture, $\hat{q} > 0$ and $\hat{p} > 0$, the seller has a unique best-response, $q^*(\hat{q},\hat{p})$, $p^*(\hat{q},\hat{p})$. This is true because (i) both $\kappa(q)$ and $c(p)$ are increasing and convex and (ii) $\frac{\partial u_S(p,q,\hat{p},\hat{q})}{\partial p}, \frac{\partial u_S(p,q,\hat{p},\hat{q})}{\partial q} > 0$, $\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial q^2}, \frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial p^2} < 0$ and $\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial p \partial q} = 0$. If the cost functions are well-behaved, then $q^*(\hat{q},\hat{p}) > 0$ is the unique solution which solves $\frac{\partial u_S(p,q,\hat{p},\hat{q})}{\partial q} = \kappa'(q)$ and $p^*(\hat{q},\hat{p}) > 0$ is the unique solution which solves $\frac{\partial u_S(p,q,\hat{p},\hat{q})}{\partial p} = c'(p)$.

Second, in equilibrium, the receiver's conjecture is correct, i.e., $\hat{q} = q^*(\hat{q},\hat{p})$ and $\hat{p} = p^*(\hat{q},\hat{p})$. This constitutes the unique informative equilibrium as long as there is a single pair of q and p (where $q = q^*(\hat{q},\hat{p}) = \hat{q} > 0$ and $p = p^*(\hat{q},\hat{p}) = \hat{p} > 0$) which solve the sender's first-order conditions. Since both $\kappa(q)$ and $c(p)$ are increasing, this implies that is sufficient to show that, *given* equilibrium behavior, the objective function is concave, i.e.,

$$\left[\frac{\partial}{\partial p}\frac{\partial u_S(p=\hat{p},q=\hat{q})}{\partial p}\right]\left[\frac{\partial}{\partial q}\frac{\partial u_S(p=\hat{p},q=\hat{q})}{\partial q}\right] > \left[\frac{\partial u_S(p=\hat{p},q=\hat{q})}{\partial p\partial q}\right]^2. \tag{34}$$

Note that

$$\left.\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial p\partial q}\right|_{\hat{p}=p,\hat{q}=q} = \frac{2p\tau(1-\omega)(q+\tau_x)}{(p(q+\tau+\tau_x)+\tau(q+\tau_x))^3}, \tag{35}$$

$$\left.\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial p^2}\right|_{\hat{p}=p,\hat{q}=q} = -\frac{2(1-\omega)(q+\tau_x)^2(q+\tau+\tau_x)}{(p(q+\tau+\tau_x)+\tau(q+\tau_x))^3} - \frac{2\omega}{(p+\tau)^3} \tag{36}$$

$$\left.\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial q^2}\right|_{\hat{p}=p,\hat{q}=q} = -\frac{2p^2(1-\omega)(p+\tau)}{(p(q+\tau+\tau_x)+\tau(q+\tau_x))^3} \tag{37}$$

which implies:

$$\Sigma \equiv \left.\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial p^2}\times\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial q^2} - \left(\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial p\partial q}\right)^2\right|_{\hat{p}=p,\hat{q}=q} \tag{38}$$

$$= -\frac{2p^2(1-\omega)\left(-\frac{2\omega(p(q+\tau+\tau_x)+\tau(q+\tau_x))^3}{(p+\tau)^2} + 2(1-\omega)(q+\tau_x)^2(1-(p+\tau)(q+\tau+\tau_x))\right)}{(p(q+\tau+\tau_x)+\tau(q+\tau_x))^6} \tag{39}$$

which implies $\Sigma > 0$ if

$$2(1-\omega)(q+\tau_x)^2(1-(p+\tau)(q+\tau+\tau_x)) < \frac{2\omega(p(q+\tau+\tau_x)+\tau(q+\tau_x))^3}{(p+\tau)^2} \tag{40}$$

$$\Leftrightarrow (1-(p+\tau)(q+\tau+\tau_x)) < \frac{2\omega(p(q+\tau+\tau_x)+\tau(q+\tau_x))^3}{2(1-\omega)(q+\tau_x)^2(p+\tau)^2} \tag{41}$$

To ensure this holds for all $p, q, \tau_x > 0$, it is sufficient to have $\tau > 1$.

Part 2 (Impact of the Curse of Knowledge)

The optimal p^* is the solution to

$$\Upsilon^p \equiv \left.\frac{\partial u_S(p,q,\hat{p},\hat{q})}{\partial p}\right|_{\hat{p}=p,\hat{q}=q} = c'(p) \tag{42}$$

and the optimal q^* is the solution to

$$\Upsilon^q \equiv \left.\frac{\partial u_S\left(p,q,\hat{p},\hat{q}\right)}{\partial q}\right|_{\hat{p}=p,\hat{q}=q} = \kappa'\left(q\right). \tag{43}$$

So differentiating both sides of each first-order condition gives:

$$\Upsilon^p_p p_\omega + \Upsilon^p_q q_\omega + \Upsilon^p_\omega = c''\left(p\right) p_\omega \tag{44}$$

and

$$\Upsilon^q_p p_\omega + \Upsilon^q_q q_\omega + \Upsilon^q_\omega = \kappa''\left(q\right) q_\omega. \tag{45}$$

Solving the system for p_ω and q_ω gives us:

$$p_\omega = \frac{\Upsilon^p_\omega\left(\kappa''(q) - \Upsilon^q_q\right) + \Upsilon^p_q \Upsilon^q_\omega}{\left(c''(p) - \Upsilon^p_p\right)\left(\kappa''(q) - \Upsilon^q_q\right) - \Upsilon^q_p \Upsilon^p_q} \tag{46}$$

$$q_\omega = \frac{\Upsilon^q_\omega\left(c''(p) - \Upsilon^p_p\right) + \Upsilon^q_p \Upsilon^p_\omega}{\left(c''(p) - \Upsilon^p_p\right)\left(\kappa''(q) - \Upsilon^q_q\right) - \Upsilon^q_p \Upsilon^p_q} \tag{47}$$

where

$$\Upsilon^p_\omega = \frac{p\tau\left(p\left(2q+\tau+2\tau_x\right)+2\tau\left(q+\tau_x\right)\right)}{(p+\tau)^2\left(p\left(q+\tau+\tau_x\right)+\tau\left(q+\tau_x\right)\right)^2} > 0 \tag{48}$$

$$\Upsilon^p_p = -\frac{\begin{pmatrix} 2p^3\left(q+\tau+\tau_x\right)\left(\tau\omega\left(2q+\tau+2\tau_x\right)+\left(q+\tau_x\right)^2\right) \\ +6p^2\tau\left(q+\tau_x\right)\left(q+\tau+\tau_x\right)\left(q+\tau\omega+\tau_x\right) \\ +6p\tau^2\left(q+\tau_x\right)^2\left(q+\tau+\tau_x\right)+2\tau^3\left(q+\tau_x\right)^2\left(q+\tau\left(1-\omega\right)+\tau_x\right) \end{pmatrix}}{(p+\tau)^3(p(q+\tau+\tau_x)+\tau(q+\tau_x))^3} < 0 \tag{49}$$

$$\Upsilon^p_q = \frac{2p\tau\left(1-\omega\right)\left(q+\tau_x\right)}{\left(p\left(q+\tau+\tau_x\right)+\tau\left(q+\tau_x\right)\right)^3} > 0 \tag{50}$$

$$\Upsilon^q_\omega = -\frac{p^2}{\left(p\left(q+\tau+\tau_x\right)+\tau\left(q+\tau_x\right)\right)^2} < 0 \tag{51}$$

$$\Upsilon^q_p = \frac{2p\tau\left(1-\omega\right)\left(q+\tau_x\right)}{\left(p\left(q+\tau+\tau_x\right)+\tau\left(q+\tau_x\right)\right)^3} > 0 \tag{52}$$

$$\Upsilon^q_q = -\frac{2p^2\left(1-\omega\right)\left(p+\tau\right)}{\left(p\left(q+\tau+\tau_x\right)+\tau\left(q+\tau_x\right)\right)^3} < 0 \tag{53}$$

Moreover, since $c\left(p\right)$ and $\kappa\left(q\right)$ are convex, we have $c''\left(p\right) > 0$ and $\kappa''\left(q\right) > 0$.

This implies

$$\left(c''(p) - \Upsilon_p^p\right)\left(\kappa''(q) - \Upsilon_q^q\right) - \Upsilon_p^q \Upsilon_q^p = \underbrace{c''(p)\kappa''(q)}_{>0} - \underbrace{c''(p)\Upsilon_q^q}_{<0} - \underbrace{\Upsilon_p^p \kappa''(q)}_{<0} + \left(\underbrace{\Upsilon_p^p \Upsilon_q^q - \Upsilon_p^q \Upsilon_q^p}_{>0}\right) \tag{54}$$

which implies:

$$\left(c''(p) - \Upsilon_p^p\right)\left(\kappa''(q) - \Upsilon_q^q\right) - \Upsilon_p^q \Upsilon_q^p > 0, \tag{55}$$

so the signs of p_ω and q_ω are determined by the numerators in the corresponding expressions. Next, note that

$$\Upsilon_\omega^p \left(\kappa''(q) - \Upsilon_q^q\right) + \Upsilon_q^p \Upsilon_\omega^q = \underbrace{\Upsilon_\omega^p \kappa''(q)}_{>0} + \underbrace{\frac{2p^3 \tau\,(1-\omega)}{(p+\tau)\left(p\left(q+\tau+\tau_x\right) + \tau\left(q+\tau_x\right)\right)^4}}_{>0} > 0,$$

which implies that $p_\omega > 0$ always. Finally, note that

$$\begin{aligned}
\Upsilon_\omega^q \left(c''(p) - \Upsilon_p^p\right) + \Upsilon_p^q \Upsilon_\omega^p &= \underbrace{\Upsilon_\omega^q c''(p)}_{<0} + \Upsilon_p^q \Upsilon_\omega^p - \Upsilon_\omega^q \Upsilon_p^p \\
&= \underbrace{\Upsilon_\omega^q c''(p)}_{<0} + \frac{2p^2\left(-\tau\omega\left(p^2(2q+\tau) + 3pq\tau + q\tau^2\right) - q(p+\tau)(pq + \tau(q-\tau)) + (p+\tau)\tau_x\left(-2p(q+\tau\omega) - (p+\tau)\tau_x + \tau(-2q - \tau\omega + \tau)\right)\right)}{(p+\tau)^3\left(p(q+\tau) + (p+\tau)\tau_x + q\tau\right)^4}
\end{aligned} \tag{56}$$

Note that the second term evaluated at $\omega = 1$ is

$$\Upsilon_p^q \Upsilon_\omega^p - \Upsilon_\omega^q \Upsilon_p^p|_{\omega=1} = -\frac{2p^2}{(p+\tau)^3\left(p(q+\tau) + (p+\tau)\tau_x + q\tau\right)^2} < 0$$

So there exists an $\bar{\omega} \in (0,1)$ such that for $\omega > \bar{\omega}$, the expression is negative. This implies that for $\omega > \bar{\omega}$, $q_\omega < 0$. □

Proof of Lemma 2

Proof. The equilibrium p^*, q^* solves

$$(1-\omega)\left(\frac{\left(\frac{1}{p} + \frac{1}{\tau_x + q}\right)^{-1}}{\tau + \left(\frac{1}{p} + \frac{1}{\tau_x + q}\right)^{-1}}\right)^2 \left(\frac{1}{q + \tau_x}\right)^2 = \kappa_q \tag{57}$$

$$\frac{1}{p^2}\left[\omega\left(\frac{\left(\frac{1}{p}\right)^{-1}}{\tau + \left(\frac{1}{p}\right)^{-1}}\right)^2 + (1-\omega)\left(\frac{\left(\frac{1}{p} + \frac{1}{\tau_x + q}\right)^{-1}}{\tau + \left(\frac{1}{p} + \frac{1}{\tau_x + q}\right)^{-1}}\right)^2\right] = c_p \tag{58}$$

Solving this system and choosing positive roots (to ensure $p, q > 0$) gives us

$$p^* = \sqrt{\frac{\omega + \kappa_q \left(\sqrt{\frac{1-\omega}{\kappa_q}} - \tau\right)^2}{c_p}} - \tau, \tag{59}$$

and

$$q^* = \left(\sqrt{\frac{1-\omega}{\kappa_q}} - \tau\right) \frac{p}{p+\tau} - \tau_x \tag{60}$$

$$= \left(\sqrt{\frac{1-\omega}{\kappa_q}} - \tau\right) \left(1 - \frac{\tau}{\sqrt{\frac{\omega+\kappa_q\left(\sqrt{\frac{1-\omega}{\kappa_q}}-\tau\right)^2}{c_p}}}\right) - \tau_x \tag{61}$$

Note that the above implies we need

$$p^* > 0 \Leftrightarrow \omega + \kappa_q \left(\sqrt{\frac{1-\omega}{\kappa_q}} - \tau\right)^2 > \tau^2 c_p \tag{62}$$

$$q^* > 0 \Leftrightarrow \frac{1-\omega}{\left(\tau + \tau_x \left(1 + \frac{\tau}{p}\right)\right)^2} > \kappa_q \tag{63}$$

The latter condition implies it is sufficient to have

$$\kappa_q < \frac{1-\omega}{\tau^2} \equiv \bar{\kappa} \tag{64}$$

and so for the first condition, it is sufficient to have

$$c_p < \frac{\omega}{\tau^2} + \frac{\kappa_q}{\tau^2} \left(\sqrt{\frac{1-\omega}{\kappa_q}} - \tau\right)^2 \equiv \bar{c}. \tag{65}$$

Next, note that

$$\frac{\partial p^*}{\partial \omega} = \frac{\tau}{2\sqrt{\frac{(1-\omega)c_p\left(1+\kappa_q\tau^2-2\tau\sqrt{(1-\omega)\kappa_q}\right)}{\kappa_q}}} > 0 \tag{66}$$

and

$$q^* = \left(\frac{1}{\psi} - \tau\right)\left(1 - \frac{\tau}{G}\right) - \tau_x \tag{67}$$

which implies that

$$\frac{\partial q^*}{\partial \psi} = -\frac{1}{\psi^2}\left(1-\frac{\tau}{G}\right)+\left(\frac{1}{\psi}-\tau\right)\frac{\tau}{G^2}\frac{\partial G}{\partial \psi} \tag{68}$$

$$= -\frac{1}{\psi^2}\left(1-\frac{\tau}{G}\right)+\left(\frac{1}{\psi}-\tau\right)\frac{\tau}{G^3}\frac{\kappa_q\tau}{\psi^2 c_p} \tag{69}$$

which is positive when

$$\left(\frac{1}{\psi}-\tau\right)\frac{\tau}{G^3}\frac{\kappa_q\tau}{c_p} > \left(1-\frac{\tau}{G}\right) \tag{70}$$

$$c_p < \left(\frac{1}{\psi}-\tau\right)\frac{\kappa_q\tau^2}{G^3\left(1-\frac{\tau}{G}\right)} \tag{71}$$

i.e., the cost of increasing p should be sufficiently low so that p increases quickly enough for the complementarity to kick in. Moreover, this is sufficient for the expected payoff to be increasing in ω, since

$$u_R = -\frac{p+q+\tau_x}{p\left(q+\tau+\tau_x\right)+\tau\left(q+\tau_x\right)}, \tag{72}$$

which is increasing in p and q. The expected payoff is increasing in Λ_s, where

$$\begin{aligned}
\Lambda_s &= \frac{q+\tau_x}{\alpha_S}\sqrt{\frac{\kappa_q}{1-\omega}} \\
&= \frac{1-\tau\sqrt{\frac{\kappa_q}{1-\omega}}}{1+\frac{\tau}{p}+\tau\Gamma} = \frac{p}{G}\left(1-\tau\sqrt{\frac{\kappa_q}{1-\omega}}\right) \\
&= \left(1-\frac{\tau}{G}\right)\left(1-\tau\sqrt{\frac{\kappa_q}{1-\omega}}\right)
\end{aligned}$$

with $\psi = \sqrt{\frac{\kappa_q}{1-\omega}}$ and $G = \frac{\sqrt{1+\kappa_q\tau^2-\frac{2\kappa_q\tau}{\psi}}}{\sqrt{c_p}}$. Note that

$$\begin{aligned}
\frac{\partial \Lambda_s}{\partial \psi} &= \left(1-\frac{\tau}{G}\right)(-\tau)+\frac{\tau}{G^2}\frac{\partial G}{\partial \psi}\left(1-\tau\psi\right) \\
&= \left(1-\frac{\tau}{G}\right)(-\tau)+\frac{\tau}{G^2}\frac{\kappa_q\tau}{c_p\psi^2 G}\left(1-\tau\psi\right) \\
&= -\tau\left(1-\frac{\tau}{G}\right)+\frac{\kappa_q\tau^2}{c_p\psi^2 G^3}\left(1-\tau\psi\right)
\end{aligned}$$

So, the curse of knowledge increases the expected payoff if

$$\frac{\kappa_q\tau}{c_p\psi^2 G^3}\left(1-\tau\psi\right) > 1-\frac{\tau}{G}$$

For ω high enough, the above condition doesn't hold and the curse of knowledge reduces the expected payoff. $\square$

B Online Appendix

B.1 Receiver Expertise and Diversity

In what follows, we generalize our analysis and endow both the sender and receiver with information about the required context, x. Specifically, suppose that context is now given by

$$x = \alpha_S x_S + \alpha_R x_R + \alpha_J x_J + \sqrt{1 - \alpha_S^2 - \alpha_R^2 - \alpha_J^2}\,\xi, \tag{73}$$

where (i) $x_S, x_R, x_J, \xi \sim N(0, 1/\tau_x)$ and mutually independent while (ii) $\alpha_S, \alpha_R, \alpha_J \in (0,1)$ and $\alpha_S^2 + \alpha_R^2 + \alpha_J^2 \leq 1$. The sender (receiver) observes x_S (x_R); both players observe x_J, and neither observes ξ.[12] Given the receiver's information, we modify the structure of the sender's communication so that

$$y = x_S + \eta. \tag{74}$$

All other features of the model remain unchanged.

The receiver's optimal action is still given by $k_R^* = \mathbb{E}_R[\theta]$ but, given his information set,

$$k_R^* = \mathbb{E}[\theta | s, y, x_R, x_J] = \Lambda(s - \mu_x(y) - \alpha_R x_R - \alpha_J x_J), \tag{75}$$

where

$$\mu_x(y) \equiv \mathbb{E}[x | y, x_R, x_J] = \alpha_S \frac{\hat{q}}{\tau_x + \hat{q}} y \tag{76}$$

$$\Lambda \equiv \frac{\tau_s}{\tau + \tau_s}, \quad \text{and} \quad \frac{1}{\tau_s} \equiv \frac{1}{\hat{p}} + \frac{\alpha_S^2}{\tau_x + \hat{q}} + \frac{1 - \alpha_S^2 - \alpha_R^2 - \alpha_J^2}{\tau_x} \tag{77}$$

and where $\hat{p}$ and $\hat{q}$ reflect the receiver's inference about the sender's choices. However, given that the sender suffers from the curse of knowledge, her inference is given by

$$\hat{k}_R = \mathbb{E}[\theta | s, y, z] = \begin{cases} \Lambda(s - \mu_x(y) - \alpha_R x_R - \alpha_J x_J) & \text{with prob } 1 - \omega \\ \Lambda_\omega(s - \alpha_S x_S - \alpha_R x_R - \alpha_J x_J) & \text{with prob } \omega \end{cases} \tag{78}$$

where

$$\Lambda_\omega = \frac{\tau_\omega}{\tau + \tau_\omega}, \quad \text{and} \quad \frac{1}{\tau_\omega} = \frac{1}{\hat{p}} + \frac{1 - \alpha_S^2 - \alpha_R^2 - \alpha_J^2}{\tau_x}. \tag{79}$$

[12] Our benchmark model is equivalent to the generalized setting in which $\alpha_S = 1$ (the sender perfectly observes the context) while $\alpha_R = \alpha_J = 0$ (the receiver was uninformed, absent any communication from the sender.)

The sender optimally chooses both q and p to maximize the objective

$$\underbrace{\mathbb{E}_S\left[-\left(\theta-\hat{k}_R\right)^2\right]}_{\equiv u_S(p,q)}-c\left(p\right)-\kappa\left(q\right). \tag{80}$$

$$=\mathbb{E}\left[\begin{array}{c}-\left(1-\omega\right)\left(\theta-\Lambda\left(s-\mu_x\left(y\right)-\alpha_R x_R-\alpha_J x_J\right)\right)^2\\ -\omega\left(\theta-\Lambda_\omega\left(s-\alpha_S x_S-\alpha_R x_R-\alpha_J x_J\right)\right)^2\end{array}\right]-c\left(p\right)-\kappa\left(q\right) \tag{81}$$

Note that this objective is identical to the one found in the benchmark model, except that both τ_s and τ_ω now reflect the receiver's additional information, x_R and x_J. This change in the information environment affects not only the sender's incentive to exert effort but also the impact of the curse of knowledge.

We refer to $\chi_S \equiv \alpha_S^2+\alpha_J^2$ ($\chi_R \equiv \alpha_R^2+\alpha_J^2$) as the sender's (receiver's) **expertise**: χ_S and χ_R measure the fraction of the context each player understands. Note that their information is potentially correlated through x_J. Holding fixed χ_R and χ_S, we refer to a decrease in α_J (i.e., a decline in the relative importance of their shared information) as an increase in the **diversity** of their expertise. To understand its impact, it is helpful to rewrite (73) as

$$x=\sqrt{\chi_S-\tfrac{1}{\delta}}x_S+\sqrt{\chi_R-\tfrac{1}{\delta}}x_R+\sqrt{\tfrac{1}{\delta}}x_J+\sqrt{1-\chi_S-\chi_R+\tfrac{1}{\delta}}\,\xi, \tag{82}$$

where δ is a proxy for the relative diversity of their information, holding fixed the players' expertise, χ_S and χ_R. In particular, note that as δ increases, the correlation between the sender's and receiver's private information falls.

We first establish the impact of expertise and diversity on the the equilibrium level of information production and communication quality.

Proposition 4. *In any informative equilibrium, endogenous communication quality, endogenous information production and, hence, the expected payoff are increasing in the players' expertise and diversity, i.e., q^* and p^* are increasing in $\alpha_S, \alpha_R, \alpha_J$, and δ.*

When her expertise (α_s) is high, the sender's understanding of the context, x_s, is more valuable in interpreting the commonly-observed signal, s. Sharing such information has a larger impact on the quality of the receiver's forecast and so, all else equal, sender expertise increases the value of communication and, hence, q^*. Similarly, as the sender's expertise increases, the receiver places more weight on the report (both Λ and Λ_ω increase) which increases the sender's perception of the value of information production and, hence p^*. Thus, in any informative equilibrium, the expected payoff increases with the sender's expertise. When the receiver has more information, he places more weight on both the sender's message as well as the report: he faces less uncertainty regarding the context and so rationally acts more aggressively to the information he observes. This increases the sender's value of reducing the noise in either signal, and so increases in either α_R or α_J lead to an increase in communication quality, information production and the expected payoff.

Holding fixed their expertise, increasing the diversity of their information implies that the sender's information is incrementally more informative for the receiver. This increases the value of communicating such information (which leads to a higher q^*) and increases the weight he places on the sender's report (which leads to a higher p^*). Together, this implies that the expected payoff increases, all else equal.

These results imply, however, that expertise and diversity also inflate the value of the information the sender *mistakenly* believes the receiver has inferred. As we show in the proof of the following lemma, this means that expertise and diversity magnify the impact of the curse of knowledge.

Lemma 3. *If either $p > 0$ or $q > 0$ is exogenous, the marginal impact of the curse of knowledge on the sender's optimal effort is increasing in the players' expertise and the diversity of their information, i.e., $\frac{\partial^3 u_S}{\partial q \partial \omega \partial \pi} < 0, \frac{\partial^3 u_S}{\partial p \partial \omega \partial \pi} > 0$ for $\pi \in \{\alpha_S, \alpha_R, \alpha_J, \delta\}$.*

The impact of the curse of knowledge on the value of communication is proportional to the weight, Λ, the receiver places on the sender's report:

$$\frac{\partial^2 u_S}{\partial q \partial \omega} = -\Lambda^2 \left(\frac{\hat{q}}{\tau_x + \hat{q}} \right)^2 \left(\frac{\alpha_s^2}{q^2} \right) \tag{83}$$

As a result, the decline in communication due to the curse of knowledge is larger when expertise (and hence the weight placed on the report, Λ) is larger.

On the other hand, the impact of the curse of knowledge on the marginal benefit of information production also depends upon the *perceived* weight, Λ_ω:

$$\frac{\partial^2 u_S}{\partial p \partial \omega} = \left[\Lambda_\omega^2 - \Lambda^2 \right] \left(\frac{1}{p^2} \right) \tag{84}$$

From equations (77) and (79), an increase in expertise has a larger effect on τ_ω than τ_s: the precision is convex in the receiver's uncertainty. We show that this also implies that expertise has a larger impact on $\Lambda_\omega = \frac{\tau_\omega}{\tau + \tau_\omega}$ than $\Lambda = \frac{\tau_s}{\tau + \tau_s}$. As a result, increases in expertise amplify the increase in information production driven by the curse of knowledge.

As with expertise, an increase in diversity increases the weight (Λ) the receiver places on the sender's report and so the value of communicating falls faster with the curse of knowledge when δ is higher. On the other hand, because the sender believes that the information intuited by the receiver is relatively more valuable when diversity is higher, i.e., $\Lambda_\omega - \Lambda$ is increasing in δ, the increase in information production due to the curse of knowledge is larger when δ is higher.

To illustrate how these two countervailing effects interact when both p and q are endogenous, we return to the linear cost setting analyzed in the paper.

Lemma 4. *With linear cost, the optimal p^* and q^* are given by*

$$p^* = \frac{G - \tau}{1 + \tau\Gamma} \tag{85}$$

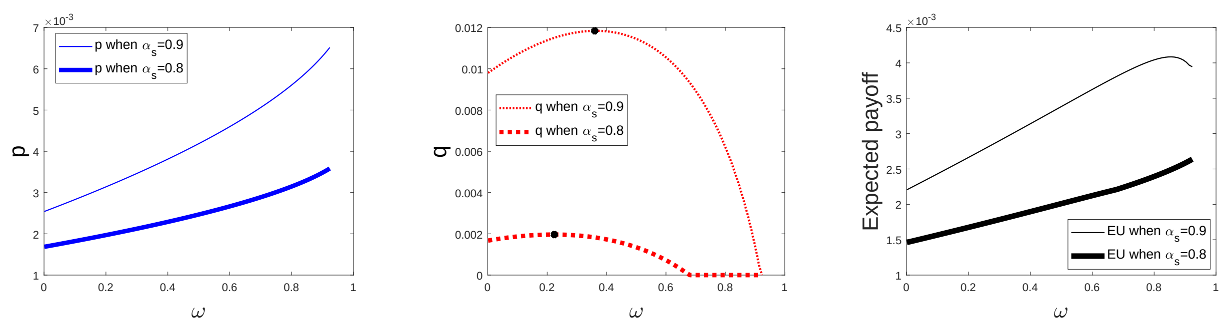
$$q^* = \frac{\alpha_S}{1 + \tau\Gamma}\left(\frac{1}{\psi} - \tau\alpha_S\right)\left(1 - \frac{\tau}{G}\right) - \tau_x \tag{86}$$

where $\Gamma \equiv \frac{1-\alpha_S^2-\alpha_R^2-\alpha_J^2}{\tau_x}$, $\psi = \sqrt{\frac{\kappa_q}{1-\omega}}$ *and* $G = \sqrt{\frac{1+\kappa_q\tau^2\alpha_S^2 - \frac{2\kappa_q\tau\alpha_S}{\psi}}{c_p}}$.
(i) The precision of the report, p^*, is always increasing in the curse of knowledge;
(ii) the quality of communication, q^*, is increasing in the curse of knowledge if

$$\frac{\kappa_q\tau^2\alpha_S}{c_p\psi G^3}\left(1 - \tau\psi\alpha_S\right) > 1 - \frac{\tau}{G} \tag{87}$$

As in the benchmark model, information production is always increasing in ω while the quality of communication can be increasing or decreasing when the receiver also privately observes some fraction of the context.

Figure 2: Optimal communication and information production with expertise
The figure plots the choice of message precision q (dashed), information production p (solid), and the expected payoff $\mathbb{E}\left[V\left(R,k\right)\right]$ under costly communication, as a function of the curse of knowledge for different values of α_S. The other parameters of the model are $\tau = 1, \tau_x = 0.01$, $\kappa_q = 0.01$ and $c_p = 0.75$

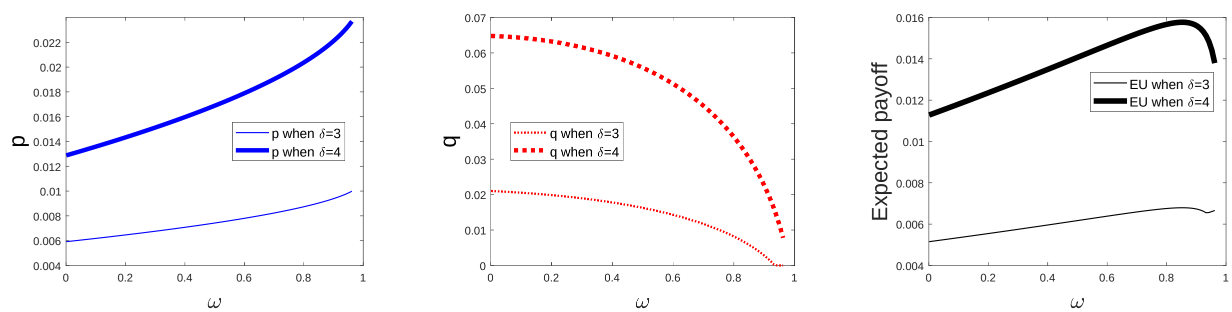


(a) p vs ω (b) q vs ω (c) Expected payoff vs ω

In Figure 2, panels (a) and (b), we plot the optimal level of information production and communication utilizing the same parameters as in Figure 1 but with different levels of expertise. Due to the linear cost function, the benefit of communication falls below the (constant) marginal cost when ω is sufficiently high (so that $q^* = 0$) which is why the informative equilibrium and, hence, the plotted lines end when $\omega < 1$. As in the benchmark model, production (p) is always increasing in ω; moreover, $\frac{dp}{d\omega}$ is increasing in the value of the sender's private information (α_S), as this increases the impact of the curse of knowledge. In panel (b), we plot the optimal level of communication as a function of the curse of knowledge. As in the benchmark example, when ω is low, the indirect effect of higher information production on q is more important and so $\frac{\partial q^*}{\partial \omega}$ is increasing in expertise. However, there exists a threshold ω (indicated by the black dot) such that communication quality begins to fall. This threshold, however, is increasing in the sender's expertise because the magnitude of the indirect effect is larger, i.e., because $\frac{dp}{d\omega}$ increases more quickly.[13] Furthermore, the increased sensitivity of both p and q to ω explains why the expected payoff increases more quickly in panel (c) when ω is sufficiently low. We find that similar patterns arise when the diversity in their information is varied, as shown in Figure 3.

Figure 3: Optimal communication and information acquisition with diversity

The figure plots the choice of message precision q (dashed), acquired information precision p (solid), and the expected payoff $\mathbb{E}[V(R,k)]$ under costly communication, as a function of the degree of cursedness for different values of δ. The other parameters of the model are set to: $\tau = 1, \tau_x = 0.01$, $\chi_S = 0.7$, $\chi_R = 0.5$, $\kappa_q = 0.01$ and $c_p = 0.75$.



(a) p vs ω (b) q vs ω (c) Expected payoff vs ω

B.2 Complementarity vs Substituitability

We adopt the generalized setting found in B.1 with one modification: the receiver's objective function is to choose his action, k_R, to minimize

$$\mathbb{E}_S\left[-\left(\beta\theta + (1-\beta)\,x - k_R\right)^2\right] \tag{88}$$

[13] We show in proof of Lemma 4 that the sign of $\frac{\partial q^*}{\partial \omega}$ does not depend on either α_R or α_J in the linear cost specification.

given his information set. Note that, given this modified objective, knowing the context isn't just important for inferring θ from the report, but it is also part of the receiver's objective. As a result, the sender's communication and the report can be substitutes.

The sender optimally chooses the quality of communication q and the precision of information produced p to maximize the objective

$$\underbrace{\mathbb{E}_S\left[-\left(\beta\theta+(1-\beta)\,x-\hat{k}_R\right)^2\right]}_{\equiv u_S(p,q,\hat{p},\hat{q})}-c\,(p)-\kappa\,(q)$$

where given the sender's curse of knowledge, her inference about the receiver's action is given by

$$\hat{k}_R=\mathbb{E}\left[x|s,y,z\right]=\begin{cases}\hat{\Lambda}_y s+\left(1-\beta-\hat{\Lambda}_y\right)\left(\hat{\Lambda}_q\alpha_S y+\alpha_R x_R+\alpha_J x_J\right) & \text{with prob } 1-\omega\\ \hat{\Lambda}_\omega s+\left(1-\beta-\hat{\Lambda}_\omega\right)\left(\alpha_S x_S+\alpha_R x_R+\alpha_J x_J\right) & \text{with prob } \omega\end{cases},\tag{89}$$

where $\hat{\Lambda}_q\equiv\frac{\hat{q}}{\hat{q}+\tau_x}$, $\hat{\Lambda}_y\equiv\hat{p}\frac{\beta\hat{\tau}_y+(1-\beta)\tau}{\hat{p}\tau+\hat{p}\hat{\tau}_y+\tau\hat{\tau}_y}$, $\hat{\Lambda}_\omega=\hat{p}\frac{\beta\tau_\omega+(1-\beta)\tau}{\hat{p}\tau+\hat{p}\tau_\omega+\tau\tau_\omega}$, $\tau_\omega=\tau_x\Gamma^{-2}$ and $\hat{\tau}_y=\frac{(\hat{q}+\tau_x)\tau_x}{\alpha_S^2\tau_x+\Gamma^2(q+\tau_x)}$

Proposition 5. *The marginal benefit of higher quality communication is increasing in the receiver's conjectured precision of the sender's report (i.e.,* $\frac{\partial^2 u_S(p,q,\hat{p},\hat{q})}{\partial q\partial\hat{p}}>0$*) if and only if* $\frac{\hat{p}}{\hat{p}+\tau}>\frac{1-\beta}{\beta}$.

In the benchmark model, communication quality can be increasing in the degree to which the sender suffers from the curse of knowledge through the bias' positive impact on information production. In particular, as we highlight in Lemma 1, when the receiver conjectures that the report is more precise (as is the case when ω increases), the value of increasing the quality of communication increases. Thus, Proposition 5 implies that, for β sufficiently high, our main mechanism still arises; however, when β is sufficiently low, the inequality in Proposition 5 cannot hold for any well-behaved cost function. For instance, if $\beta=0$, θ is simply noise and s and y are signals of x, i.e., the report and the sender's communication are substitutes given the objective function.

B.3 Robustness

B.3.1 Overt communication: receiver observes both p and q

Given the receiver's objective function, his optimal action is given by

$$k_R=\mathbb{E}_R\left[\theta\right]=\mathbb{E}\left[\theta|s,y\right].\tag{90}$$

Using standard properties of normal distributions, we can express this as

$$k_R = \Lambda\left(s - \mu_x(y)\right), \text{ where } \mu_x(y) \equiv \mathbb{E}[x|y] = \lambda y \tag{91}$$

and where

$$\lambda \equiv \frac{q}{\tau_x + q}, \quad \Lambda \equiv \frac{\tau_s}{\tau + \tau_s}, \quad \text{and} \quad \frac{1}{\tau_s} \equiv \frac{1}{p} + \frac{1}{\tau_x + q}. \tag{92}$$

However, given the sender's curse of knowledge, her inference about the receiver's action is given by

$$\hat{k}_R = \mathbb{E}[\theta|s, y, z] = \begin{cases} \Lambda\left(s - \mu_x(y)\right) & \text{with prob } 1-\omega \\ \Lambda_\omega\left(s - x\right) & \text{with prob } \omega \end{cases}, \tag{93}$$

where

$$\Lambda_\omega = \frac{\tau_\omega}{\tau + \tau_\omega}, \quad \text{and} \quad \frac{1}{\tau_\omega} = \frac{1}{p}. \tag{94}$$

Given her beliefs about the receiver's action, the sender optimally chooses the quality of communication q and the precision of information produced p to maximize the objective

$$\begin{aligned} u_S(p, q, \hat{p}, \hat{q}) &\equiv \mathbb{E}_S\left[-\left(\theta - \hat{k}_R\right)^2\right] \\ &= -\mathbb{E}\left[\ (1-\omega)\left(\theta - \Lambda\left(s - \mu_x(y)\right)\right)^2 \ + \omega\left(\theta - \Lambda_\omega\left(s - x\right)\right)^2\right] \end{aligned}$$

subject to the cost $c(p) + \kappa(q)$. The sender's marginal benefit of increasing her effort along either dimension can be written as:

$$\frac{\partial}{\partial p} u_S(p, q, \hat{p}, \hat{q}) = \frac{1}{p^2}\left(\omega \Lambda_\omega^2 + (1-\omega)\Lambda^2\right) \tag{95}$$

and

$$\frac{\partial}{\partial q} u_S(p, q, \hat{p}, \hat{q}) = (1-\omega)\Lambda^2 \frac{\lambda^2}{q^2}. \tag{96}$$

These first-order conditions are identical to the ones found in the baseline model, which implies that the equilibrium precision choices are unchanged.

B.3.2 Sender observes x before choosing p, q

The sender optimally chooses the quality of communication q and the precision of information produced p to maximize the objective

$$u_S(p, q, x) \equiv \mathbb{E}_S\left[-\left(\theta - \hat{k}_R\right)^2\right] - c(p) - \kappa(q). \tag{97}$$

$$= \mathbb{E}\begin{bmatrix} -(1-\omega)(\theta - \Lambda(s - \mu_x(y)))^2 \\ -\omega(\theta - \Lambda_\omega(s - x))^2 \end{bmatrix} - c(p) - \kappa(q) \tag{98}$$

We can simplify each of the expectations as

$$\mathbb{E}\left[(\theta - \Lambda(s - \mu_x(y)))^2 | x\right] = \frac{\tau}{(\tau + \tau_s)^2} + \left(\frac{\tau_s}{\tau + \tau_s}\right)^2 \left[\left(\frac{\tau_x}{\tau_x + \hat{q}}\right)^2 x^2 + \left(\frac{\hat{q}}{\tau_x + \hat{q}}\right)^2 \tfrac{1}{q} + \tfrac{1}{p}\right]$$

$$\mathbb{E}\left[(\theta - \Lambda_\omega(s - \alpha_S x_S))^2 | x_s\right] = \frac{\tau}{(\tau + \tau_\omega)^2} + \left(\frac{\tau_\omega}{\tau + \tau_\omega}\right)^2 \frac{1}{p}.$$

The only difference (relative to the setting in which the context x is not known ex-ante) arises in the first expectation but does not affect the first-order conditions since the sender takes $\hat{q}$ and τ_s as given when choosing what to communicate. Thus, the sender's first-order conditions are identical to the expressions in the baseline model.

B.4 Proofs for the Online Appendix

Proof of Proposition 4

Proof. The sender's first-order conditions are

$$(1-\omega)\left(\frac{\tau_s}{\tau + \tau_s}\right)^2 \left(\frac{\hat{q}}{\tau_x + \hat{q}}\right)^2 \left(\tfrac{\alpha_s^2}{q^2}\right) = \kappa'(q) \tag{99}$$

$$\left[(1-\omega)\left(\frac{\tau_s}{\tau + \tau_s}\right)^2 + \omega\left(\frac{\tau_\omega}{\tau + \tau_\omega}\right)^2\right]\left(\tfrac{1}{p^2}\right) = c'(p) \tag{100}$$

The optimal p^* is the solution to

$$\Upsilon^p \equiv \left.\frac{\partial u_S(p, q, \hat{p}, \hat{q})}{\partial p}\right|_{\hat{p}=p, \hat{q}=q} = c'(p) \tag{101}$$

and the optimal q^* is the solution to

$$\Upsilon^q \equiv \left.\frac{\partial u_S(p, q, \hat{p}, \hat{q})}{\partial q}\right|_{\hat{p}=p, \hat{q}=q} \kappa'(q), \tag{102}$$

So differentiating both sides of each first order condition w.r.t. $\pi \in \{\alpha_R, \alpha_S, \alpha_J, \delta\}$ gives:

$$\Upsilon_p^p p_\pi + \Upsilon_q^p q_\pi + \Upsilon_\pi^p = c''(p)\, p_\pi \tag{103}$$

and

$$\Upsilon_p^q p_\pi + \Upsilon_q^q q_\pi + \Upsilon_\pi^q = \kappa''(q)\, q_\pi. \tag{104}$$

Solving the system for p_π and q_π gives us:

$$p_\pi = \frac{\Upsilon_\pi^p \left(\kappa''(q) - \Upsilon_q^q\right) + \Upsilon_q^p \Upsilon_\pi^q}{\left(c''(p) - \Upsilon_p^p\right)\left(\kappa''(q) - \Upsilon_q^q\right) - \Upsilon_p^q \Upsilon_q^p} \tag{105}$$

$$q_\pi = \frac{\Upsilon_\pi^q \left(c''(p) - \Upsilon_p^p\right) + \Upsilon_p^q \Upsilon_\pi^p}{\left(c''(p) - \Upsilon_p^p\right)\left(\kappa''(q) - \Upsilon_q^q\right) - \Upsilon_p^q \Upsilon_q^p} \tag{106}$$

As before, one can show that $\Upsilon_p^p, \Upsilon_q^q < 0$ and $\Upsilon_p^q > 0$ so that

$$\left(c''(p) - \Upsilon_p^p\right)\left(\kappa''(q) - \Upsilon_q^q\right) - \Upsilon_p^q \Upsilon_q^p > 0$$

Moreover, one can verify that $\Upsilon_\pi^p, \Upsilon_\pi^q > 0$ for $\pi \in \{\alpha_R, \alpha_S, \alpha_J, \delta\}$. This implies

$$\Upsilon_\pi^p \left(\kappa''(q) - \Upsilon_q^q\right) + \Upsilon_q^p \Upsilon_\pi^q > 0 \Leftrightarrow p_\pi > 0 \tag{107}$$

$$\Upsilon_\pi^q \left(c''(p) - \Upsilon_p^p\right) + \Upsilon_p^q \Upsilon_\pi^p > 0 \Leftrightarrow q_\pi > 0 \tag{108}$$

which establishes the result. □

Proof of Lemma 3

Proof. We know that

$$\frac{\partial^2 u_S}{\partial p \partial \omega} = \left[\left(\frac{\tau_\omega}{\tau + \tau_\omega}\right)^2 - \left(\frac{\tau_s}{\tau + \tau_s}\right)^2\right]\left(\tfrac{1}{p^2}\right) > 0 \tag{109}$$

and so

$$\frac{\partial^3 u_S}{\partial p \partial \omega \partial \alpha} = \left[\left(\frac{\tau_\omega}{\tau + \tau_\omega}\right)\left(\frac{\frac{\partial \tau_\omega}{\alpha}}{\tau + \tau_\omega}\right) - \left(\frac{\tau_s}{\tau + \tau_s}\right)\left(\frac{\frac{\partial \tau_s}{\alpha}}{\tau + \tau_s}\right)\right]\left(\tfrac{2\tau}{p^2}\right). \tag{110}$$

Utilizing equations (77) and (79), this implies that

$$\frac{\partial^3 u_S}{\partial p \partial \omega \partial \alpha_R} = \left[\Lambda_\omega^2 \tau_\omega - \Lambda^2 \tau_s\right] \left(\tfrac{4\alpha_R \tau}{\tau_x p^2}\right) > 0 \tag{111}$$

$$\frac{\partial^3 u_S}{\partial p \partial \omega \partial \alpha_J} = \left[\Lambda_\omega^2 \tau_\omega - \Lambda^2 \tau_s\right] \left(\tfrac{4\alpha_J \tau}{\tau_x p^2}\right) > 0 \tag{112}$$

$$\frac{\partial^3 u_S}{\partial p \partial \omega \partial \alpha_J} = \left[\Lambda_\omega^2 \frac{\tau_\omega}{\tau_x} - \Lambda^2 \tau_s \left(\frac{1}{\tau_x} - \frac{1}{\tau_x + \hat{q}}\right)\right] \left(\tfrac{4\alpha_J \tau}{\tau_x p^2}\right) > 0 \tag{113}$$

where the inequalities follow from the fact that $\Lambda_\omega > \Lambda$ and $\tau_\omega > \tau_s$.

Similarly, note that

$$\frac{\partial^2 u_S}{\partial q \partial \omega} = -\left(\frac{\tau_s}{\tau + \tau_s}\right)^2 \left(\frac{\hat{q}}{\tau_x + \hat{q}}\right)^2 \left(\tfrac{\alpha_s^2}{q^2}\right) < 0 \tag{114}$$

and so by equations (77) and (79),

$$\frac{\partial^3 u_S}{\partial q \partial \omega \partial \alpha_S} = -2\Lambda^2 \left(\frac{\hat{q}}{\tau_x + \hat{q}}\right)^2 \left[\left(\tfrac{\alpha_s}{q^2}\right) + 2\alpha_S \tau_s \tau \left(\frac{1}{\tau_x} - \frac{1}{\tau_x + \hat{q}}\right) \left(\tfrac{\alpha_s^2}{q^2}\right)\right] < 0 \tag{115}$$

$$\frac{\partial^3 u_S}{\partial q \partial \omega \partial \alpha_R} = -4\Lambda^2 \left(\frac{\alpha_R \tau \tau_s}{\tau_x}\right) \left(\frac{\hat{q}}{\tau_x + \hat{q}}\right)^2 \left(\tfrac{\alpha_s^2}{q^2}\right) < 0 \tag{116}$$

$$\frac{\partial^3 u_S}{\partial q \partial \omega \partial \alpha_J} = -4\Lambda^2 \left(\frac{\alpha_J \tau \tau_s}{\tau_x}\right) \left(\frac{\hat{q}}{\tau_x + \hat{q}}\right)^2 \left(\tfrac{\alpha_s^2}{q^2}\right) < 0. \tag{117}$$

We know that

$$\frac{\partial^3 u_S}{\partial p \partial \omega \partial \delta} = \tfrac{2\tau}{\tau_x \delta^2 p^2} \left[\Lambda_\omega^2 \tau_\omega - \Lambda^2 \tau_s \left(\frac{\hat{q}}{\tau_x + \hat{q}}\right)\right] > 0 \tag{118}$$

since $\Lambda_\omega > \Lambda$, $\tau_\omega > \tau_s$. Similarly, we know that

$$\frac{\partial^3 u_S}{\partial q \partial \omega \partial \delta} = -2\left(\frac{\hat{q}}{\tau_x + \hat{q}}\right)^2 \frac{\Lambda^2}{\delta^2} \left[\frac{\tau \tau_s \hat{q}}{\tau_x (\tau_x + \hat{q})} + \frac{1}{q^2}\right] < 0. \tag{119}$$

□

Proof of Lemma 4

Proof. The optimal p^* is the solution to

$$\frac{1}{p^2}\left[\omega \Lambda_\omega^2 + (1-\omega)\Lambda_s^2\right] = c_p$$

$$(1-\omega)\Lambda_s^2 \alpha_S^2 \left(\frac{1}{q + \tau_x}\right)^2 = \kappa_q$$

where

$$\Lambda_s \equiv \frac{\tau_s}{\tau+\tau_s}, \quad \text{and} \quad \frac{1}{\tau_s} \equiv \frac{1}{p} + \frac{\alpha_S^2}{\tau_x+q} + \frac{1-\alpha_S^2-\alpha_R^2-\alpha_J^2}{\tau_x}$$

$$\Lambda_\omega = \frac{\tau_\omega}{\tau+\tau_\omega}, \quad \text{and} \quad \frac{1}{\tau_\omega} = \frac{1}{p} + \frac{1-\alpha_S^2-\alpha_R^2-\alpha_J^2}{\tau_x}.$$

The second FOC implies

$$q+\tau_x = \frac{\alpha_S\sqrt{\frac{1-\omega}{\kappa_q}} - \tau\alpha_S^2}{1+\frac{\tau}{p}+\tau\Gamma}$$

where $\frac{1-\alpha_S^2-\alpha_R^2-\alpha_J^2}{\tau_x} \equiv \Gamma$. Solving the FOCs gives

$$p = \frac{\tau_x}{\tau\left(1-\alpha_S^2-\alpha_R^2-\alpha_J^2\right)+\tau_x}\left(\frac{\sqrt{\omega+\kappa_q\left(\sqrt{\frac{1-\omega}{\kappa_q}}-\tau\alpha_S\right)^2}}{\sqrt{c_p}}-\tau\right) \tag{120}$$

$$= \frac{G-\tau}{1+\tau\Gamma} \tag{121}$$

Substituting this into the expression for q gives us

$$q = \frac{\alpha_S\sqrt{\frac{1-\omega}{\kappa_q}} - \tau\alpha_S^2}{1+\frac{\tau}{p}+\tau\Gamma} - \tau_x \tag{122}$$

$$= \frac{\alpha_S}{1+\tau\Gamma}\left(\frac{1}{\psi}-\tau\alpha_S\right)\left(1-\frac{\tau}{G}\right)-\tau_x \tag{123}$$

This increases in ω if $\frac{\partial q}{\partial \psi} > 0$ i.e.,

$$\frac{\kappa_q\tau^2\alpha_S}{c_p\psi G^3}\left(1-\tau\psi\alpha_S\right) > 1-\frac{\tau}{G} \tag{124}$$

This implies that the sign of $\frac{\partial q^*}{\partial \omega}$ does not depend on either α_R or α_J in the linear cost specification.

□

Proof of Proposition 5

Proof. Note that the objective of the sender is

$$\mathbb{E}\left[\begin{array}{c} -(1-\omega)\left(\theta+x-\left(\hat{\Lambda}_y s+\left(1-(1-\beta)\hat{\Lambda}_y\right)\left(\hat{\Lambda}_q\alpha_S y+\alpha_R x_R+\alpha_J x_J\right)\right)\right)^2 \\ -\omega\left(\theta+x-\left(\hat{\Lambda}_\omega s+\left(1-\hat{\Lambda}_\omega(1-\beta)\right)\left(\alpha_S x_S+\alpha_R x_R+\alpha_J x_J\right)\right)\right)^2 \end{array}\right] - c(p) - \kappa(q)$$

The first term in the above expectation can be simplified as

$$\frac{\left(\beta-\hat{\Lambda}_y\right)^2}{\tau}+\left(1-\beta-\hat{\Lambda}_y\right)^2\left[\frac{\Gamma^2}{\tau_x}+\frac{\alpha_S^2\left(1-\hat{\Lambda}_q\right)^2}{\tau_x}+\frac{\alpha_S^2\hat{\Lambda}_q^2}{q}\right]+\frac{\hat{\Lambda}_y^2}{p}$$

and the second term can be simplified as

$$\frac{\left(\beta-\hat{\Lambda}_\omega\right)^2}{\tau}+\frac{\left(1-\beta-\hat{\Lambda}_\omega\right)^2\Gamma^2}{\tau_x}+\frac{\hat{\Lambda}_\omega^2}{p}$$

This implies that

$$\frac{\partial u_S\left(p,q,\hat{p},\hat{q}\right)}{\partial q}=(1-\omega)\left(1-\beta-\hat{\Lambda}_y\right)^2\hat{\Lambda}_q^2\left(\frac{\alpha_S^2}{q^2}\right)>0$$

and

$$\frac{\partial^2 u_S\left(p,q,\hat{p},\hat{q}\right)}{\partial q\partial\hat{p}}=-2\left(1-\omega\right)\left(1-\beta-\hat{\Lambda}_y\right)\hat{\Lambda}_q^2\left(\frac{\alpha_S^2}{q^2}\right)\frac{\partial\hat{\Lambda}_y}{\partial\hat{p}}$$

where $\frac{\partial\hat{\Lambda}_y}{\partial\hat{p}}=\left(\frac{\beta\tau_y+(1-\beta)\tau}{(\hat{p}\tau+\hat{p}\tau_y+\tau\tau_y)^2}\right)\tau\tau_y>0$. This implies that $\frac{\partial^2 u_S(p,q)}{\partial q\partial\hat{p}}>0$ iff $1-\beta-\hat{\Lambda}_y<0$ which is the case if and only if

$$\frac{\hat{p}}{\hat{p}+\tau}>\frac{1-\beta}{\beta}.$$

□