9

The Evolution of Deliberate Ignorance in Strategic Interaction

Christian Hilbe and Laura Schmid

Abstract

Optimal decision making requires individuals to know their available options and to anticipate correctly what consequences these options have. In many social interactions, however, we refrain from gathering all relevant information, even if this information would help us make better decisions and is costless to obtain. This chapter examines several examples of "deliberate ignorance." Two simple models are proposed to illustrate how ignorance can evolve among self-interested and payoff-maximizing individuals, and open problems are highlighted that lie ahead for future research to explore.

Introduction

Information is a precious resource. Common sense suggests that we make better decisions the more we understand our options and their consequences. Yet people sometimes deliberately choose not to gather relevant information, even if this information is readily available and effective to make better choices (see Hertwig and Engel, this volume, 2016). For instance, people give considerable amounts to charities, but they rarely consider how efficient a charity is, although there are websites that allow for easy comparison (Hoffman et al. 2016). Managers often avoid arguments that run counter to their previous decisions, although such arguments would help them to abandon projects with a low probability of success in a timely fashion (Deshpandé and Kohli 1989). When asked for a costly favor, subjects in laboratory experiments sometimes avoid retrieving information about the exact cost, especially if information retrieval is observable by outside parties (Jordan et al. 2016). These and other instances of ignorance can be the result of different cognitive processes. Individuals may physically avoid learning a piece of evidence, they may stop paying attention, they may deliberately misconstrue evidence, or they may try

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

selectively to forget information that is considered unpleasant (Golman et al. 2017). Here we wish to sketch how mathematical models can illuminate such paradoxical behaviors and discuss why some of these behaviors seem more puzzling than others.

In the following, when we refer to "ignorance," we mean that an individual does not know of a certain fact which, if known, could affect some of the individual's future decisions. We say this ignorance is "deliberate" if the individual had a chance to learn this fact but chose not to. In particular, deliberate ignorance requires individuals to be aware of gaps in their knowledge, and they must have the means to resolve these gaps. However, we do not require deliberate ignorance to be the result of a calculated process in which individuals consciously weigh the advantages and disadvantages of further information. Instead, we wish to emphasize that deliberate ignorance can endogenously evolve among individuals who repeatedly encounter similar decision problems, and who adapt their strategies based on simple heuristics.

Some cases of deliberate ignorance are amenable to a straightforward economic explanation. In many cases, individuals simply avoid information because it seems irrelevant or because its expected benefits do not warrant the search costs (Stigler 1961). For example, many of us will not know how to react properly if we are attacked by a bear, presumably because such an event appears too unlikely to justify even one minute of Internet search.[1] Other instances of deliberate ignorance, in which information is essentially costless, are subtler to fathom. In addressing these more interesting cases, we find it useful to distinguish between strategic and nonstrategic ignorance, which is a somewhat coarser distinction than the one proposed by Hertwig and Engel (this volume, 2016). In models of strategic ignorance, information is typically taken as a means to an end. It has no intrinsic value, but it allows individuals to evaluate their options better. Instances of strategic ignorance include cases in which individuals avoid information in order to commit themselves credibly to a certain path of action (Schelling 1960), when they want to avoid leaking information or biasing themselves in a negotiation (Auster and Dana, this volume), or when they exploit a moral wiggle room when making morally ambiguous decisions (Dana et al. 2007). In models of nonstrategic ignorance, the mere possession of information (or the way by which it was obtained) may affect a person's well-being. For example, people might avoid information due to regret aversion or dissonance avoidance, even if the information itself would allow them to take actions that would improve their future material payoffs.

In the following two sections, we introduce two simple models that illustrate some of the issues that arise when modeling the evolution of strategic

[1] It turns out that the proper reaction depends on the species of bear. The U.S. National Park Service advises to play dead when being attacked by a grizzly, yet to escape to a secure place when being attacked by a black bear; see https://www.nps.gov/subjects/bears/safety.htm (accessed May 8, 2019).

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

ignorance. In the final section, we briefly discuss models of nonstrategic ignorance, which are somewhat more difficult to grasp from an evolutionary perspective.

Ignorance as a Commitment Device

As popularized by Schelling (1960), players can use self-commitment as a powerful tool to improve their strategic position. The idea is that by eliminating some of their strategic options, players can enhance the credibility of pledges that would otherwise be viewed simply as cheap talk. Self-commitment can take various forms, ranging from the proverbial burning of bridges to disabling one's steering wheel in the game of chicken. As noted by Schelling, avoiding certain kinds of information can act as a commitment device as well, such as when second movers deliberately ignore the action of the first mover.

To illustrate the value of deliberate ignorance as a commitment device, let us consider the "envelope game" (Bear and Rand 2019; Hilbe et al. 2015; Hoffman et al. 2015). The envelope game is a stylized model used to illustrate the tensions that arise when players cooperate for opportunistic reasons (when cooperation happens to be to their own advantage) or out of principle (no matter how the current incentives for cooperation are). The game involves two players and has four consecutive stages (Figure 9.1). In the first stage, a chance move by nature (N) determines whether the players face an environment in which cooperation carries a high (H) or a low (L) cost. The outcome of this chance move cannot be observed directly; the players only know that on average they face a high-cost environment with probability p. In the second stage, player 1 has the option to learn the current state of the environment at no cost. Based on player 1's decision, player 2 can choose whether to accept or reject the pairwise interaction in the third stage. If the interaction is rejected, the game is over and both players receive a default payoff of zero. Otherwise, if the interaction is accepted, the game enters a fourth stage in which player 1 chooses whether or not to cooperate with player 2. If player 1 cooperates, each player i receives a benefit $b_i > 0$. However, player 1 also needs to bear a cost c_s. This cost depends on the current state, $s \in \{L, H\}$, of the environment with $c_L < c_H$. If player 1 defects, each player i receives a payoff of d_i.

For the game to be interesting, we assume that payoffs satisfy the following two conditions. First, player 2 prefers cooperative interactions to no interaction, but strongly opposes interactions with a defector:

$$b_2 > 0 > d_2. \tag{9.1}$$

This condition can be seen as a definition of what it means to "cooperate": taking an action that is to the co-player's advantage even if it may be individually costly. Second, P1 prefers to cooperate only in a low-cost environment:

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

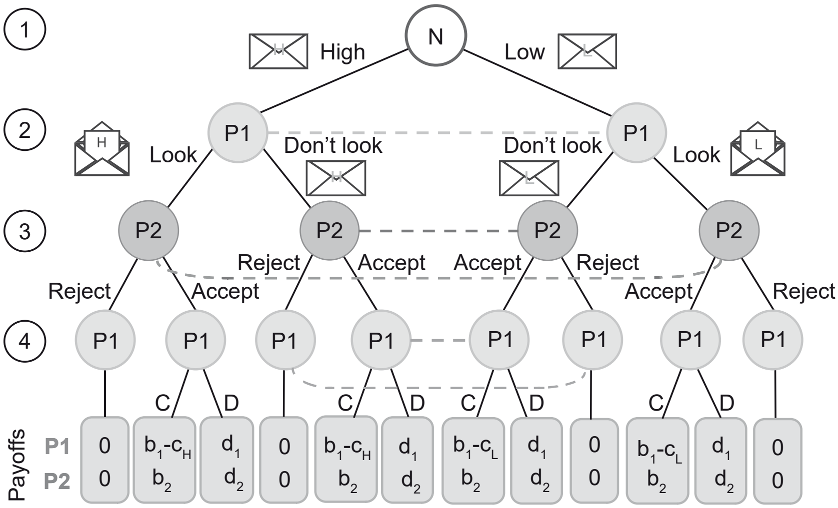

Figure 9.1 The envelope game is an asymmetric game with incomplete information between two players, P1 and P2, and four stages: (1) Nature (N) determines randomly whether players are in a high-cost (H) or low-cost environment (L). Neither of the two players knows the state of the environment (as illustrated by the closed envelope). (2) P1 decides whether or not to look into the envelope to learn the state of the environment. (3) Based on whether the envelope has been opened, P2 chooses whether or not to accept P1 as an interaction partner. If P2 rejects P1, the game is over and both players receive no payoff. (4) If accepted, P1 decides whether to cooperate or defect. If P1 has opened the envelope in stage 2, this decision may be contingent on the realized cooperation cost. Payoffs are such that P2 always prefers P1 to cooperate. However, P1 only has an incentive to do so in a low-cost environment. Dashed lines represent information sets, connecting nodes that the respective players cannot distinguish, given the information they have.

$$b_1 - c_L > d_1 > b_1 - c_H. \tag{9.2}$$

This condition guarantees that the information in the second stage is useful in subsequent stages. Alone, player 1 would prefer to learn the state of the environment and to cooperate conditionally.

This envelope game can be solved by backward induction (see Appendix 9.1 for details). Depending on the probability p of a high-cost environment, there are three possible outcomes. First, if p is comparably small, the players are predicted to settle at an "opportunism equilibrium." In this equilibrium, player 1 learns the state of the environment, player 2 accepts the interaction, and player 1 cooperates whenever own interest is best served (i.e., only if the cost is low). The rationale for this equilibrium is straightforward: As long as high-cost environments are rare, player 2 accepts co-players who occasionally find it worthwhile to defect.

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

Second, if the high-cost probability p is very large, player 2 will always reject the interaction, independent of whether or not player 1 chose to learn the current environment. Again, this "no interaction equilibrium" is straightforward to rationalize. If the cost of cooperation is typically high, player 2 either expects player 1 to defect by default (if player 1 does not know the current environment) or is sufficiently likely to defect (if player 1 learned the environmental state in stage 2). In between these two extremes,

$$\frac{b_2}{b_2 - d_2} < p < \frac{b_1 - d_1 - c_L}{c_H - c_L}, \tag{9.3}$$

there is an "ignorance equilibrium." In this equilibrium, player 1 deliberately ignores the current state of the environment in the second stage. This leads player 2 to accept the interaction in the third stage, and player 1 cooperates in the final stage. The second inequality in Equation 9.3 ensures that high-cost environments are sufficiently rare such that player 1 cooperates by default. At the same time, the first inequality in Equation 9.3 ensures that high-cost environments are too common (or too harmful) for player 2 to accept purely opportunistic co-players.

For these predictions to be sensible, we do not need to require that players derive their strategies through rational calculation. Instead, it suffices that individuals adapt their strategies over time based on the past success that they have had. To illustrate this point, Figure 9.2 shows the dynamics of the envelope game in populations of players who adopt new strategies by imitating peers with a higher payoff (Traulsen and Hauert 2009; see Appendix 9.1 for the exact setup). These simulations recover the previously predicted equilibrium outcomes. In particular, for intermediate values of the probability p, we observe that subjects in the role of player 1 learn to ignore the costs of cooperation and they tend to cooperate unconditionally. In the end, players act *as if* they performed Bayesian updating and backward induction although they never make the respective calculations. These results allow several observations:

When ignorance pays. According to our results, deliberate ignorance is most likely to emerge when p is intermediate; that is, when the hidden information would actually be most valuable to player 1. This finding underlines the function of ignorance as a commitment device. It allows player 1 to persuade others to engage in an interaction that they otherwise would be reluctant to accept. In addition, the inequalities (Equation 9.3) suggest that strategic ignorance is most likely to emerge if defection is very costly to player 2 (i.e., if the value of d_2 is small) and if cooperation is relatively cheap for player 1 even in a high-cost environment (i.e., if $b_1 - c_H$ and d_1 are of similar magnitude).

Ignorance and altruism. In the ignorance equilibrium, player 1 will sometimes cooperate although his immediate material incentives happen to make cooperation unprofitable. To an outside observer who only observes the final stage of the game and the resulting payoffs, these instances of cooperation will

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

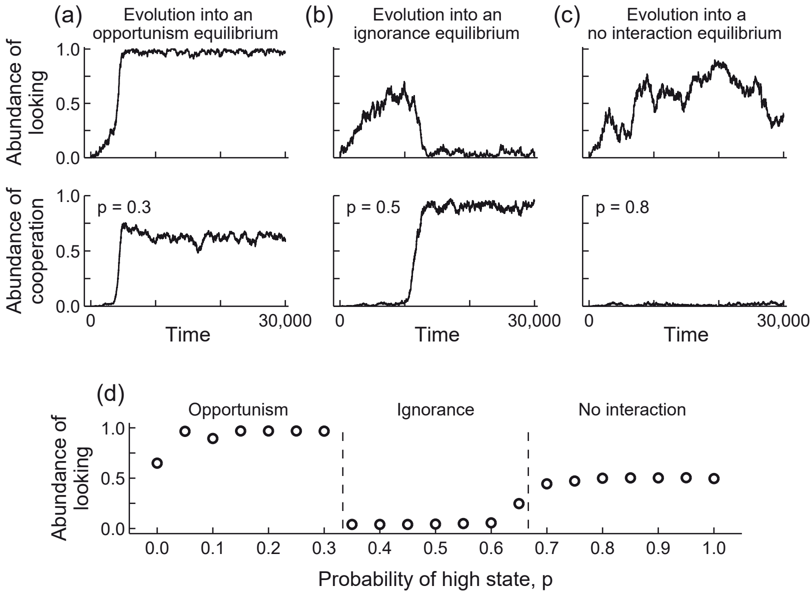

Figure 9.2 Simulating the evolution of ignorance in the envelope game, where there are two distinct populations. Members of population 1 are randomly matched with members of population 2 to interact in the envelope game (members of population *i* act in the role of player *i*). Each player is equipped with a strategy. The strategies tell the players what to do in each stage in which they need to make a decision. After these interactions, strategies with a high payoff reproduce within the respective population, either because successful players have more offspring (genetic inheritance) or because they are imitated more often (cultural inheritance). Panels (a)–(c) show three representative trajectories of this evolutionary process. If high-cost environments are sufficiently rare (i.e., if *p* is sufficiently small), players in population 1 tend to open the envelope and cooperate if the cost is low. If there is an intermediate risk of a high-cost environment, player 1 cooperates without looking. Finally, if costs are likely to be high, player 2 typically rejects all interactions no matter whether player 1 looked. In panel (d), black dots show time averages of simulation runs for different probabilities of a high-risk environment. The dashed vertical lines in panel D indicate the three different equilibrium outcomes according to backward induction.

appear as if player 1 acts altruistically. Altruistic cooperation has been observed in numerous experiments on behavior in social dilemmas. Subjects sometimes cooperate even if interactions are one-shot and fully anonymous (Dawes et al. 2007; Fehr and Fischbacher 2003). To explain these apparently irrational instances of cooperation, researchers typically argue that subjects have social preferences (Bolton and Ockenfels 2000; Fehr and Schmidt 1999), or that subjects make their decisions based on simple heuristics (Bear and Rand 2016; Delton et al. 2011; Jagau and van Veelen 2017). These heuristics are thought to be generally adaptive in the players' natural environment even if they may misfire in exceptional cases (Fawcett et al. 2014; Gigerenzer and Goldstein 1996; Hertwig et al. 2013). The above model gives an alternative interpretation

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

for why social heuristics may have evolved. If intuitive cooperators are considered more reliable, people may, in turn, have an incentive to avoid learning the details of a strategic interaction and to cooperate instinctively (Hoffman et al. 2016). In line with this argument, laboratory experiments suggest that subjects tend to be more cooperative when they need to make quick decisions (Rand et al. 2012), and uncalculating cooperators are considered more trustworthy (Jordan et al. 2016). Conversely, Ma et al. (2018) find that trustees in a trust game tend to be more generous toward those investors who were less inquisitive about the trustee's past history.

Observable ignorance. In the above model, we have assumed that player 2 observes whether or not player 1 decided to learn the state of the environment. This observability is crucial for our results (in fact, for all self-commitment models). If there was an option to secretly learn the current state, it would be a weakly dominant action for player 1 to do so. In equilibrium, player 2 would expect player 1 to know the environmental state and, as a consequence, would reject all interactions (provided the first inequality in Equation 9.3 holds). In particular, although people can commit themselves by publicly refusing to learn some relevant information, they cannot do so by engaging in similar acts of internal self-commitment, such as forgetting relevant information or privately misconstruing it.

Ignorance in the presence of communication. For the standard version of the envelope game depicted in Figure 9.1, we have assumed that the players cannot directly communicate with each other. In particular, we have assumed that player 1 is unable to inform player 2 about the present state of the environment after opening the envelope. If such communication was possible, the equilibrium predictions may change, depending on the defection payoff d_1 of player 1. If $d_1 < 0$, player 1 prefers not to interact with player 2, rather than to defect. Thus, when player 1 finds out that the present costs of cooperation are high, it is in his own interest to communicate this fact truthfully to his co-player, such that player 2 can abort the interaction. If communication is possible, there is thus only one equilibrium when $d_1 < 0$. In this equilibrium, player 1 always looks, communicates the result to player 2, and player 2 reacts accordingly. In contrast, communication has no effect when $d_1 > 0$. In that case, player 1 always prefers to interact. As a result, even when players are in a high-cost environment, player 1 has an incentive to pretend the cooperation cost is low. Thus, any message of player 1 represents cheap talk. Knowing this, player 2 does best by ignoring all communication. Consequently, we yield the same equilibria as in the no-communication case.

Ignorance versus deliberate ignorance. For cooperation to evolve under the conditions of Equation 9.3, it is actually not necessary that player 1 actively decides to avoid information. Instead, we yield the same cooperation rates if player 1 were not even given the option to learn the environmental state (i.e., if stage 2 was removed from the game altogether). That is, for intermediate values of p, the evolution of cooperation only requires *ignorance*, not *deliberate*

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

ignorance. In the above model, deliberate ignorance only emerges as a by-product, as a means to ensure that player 1 remains uninformed. In contrast, in the following we present a variation of the envelope game in which the active choice not to know is crucial.

Ignorance as a Costly Signal

The following model variation is based on the idea that different types of player 1 may have different incentives to act opportunistically (Pérez-Escudero et al. 2016). In that case, the decision of player 1 to ignore relevant information may not only communicate the player's commitment but also the player's type.

To incorporate this idea, we introduce an additional stage to the envelope game. In this stage 0, nature randomly determines the type of player 1. We assume that with probability q, player 1 is "unfavorable" (U), whereas with probability $1 - q$, player 1 is "favorable" (F). Player 1 always knows his own type, but player 2 only knows the general probability q. The two types differ in their respective likelihood to face a high-cost environment and in their incentives to cooperate. Specifically, we assume that while favorable players encounter a high-cost environment in stage 1 with probability p_F, the respective probability for unfavorable players is $p_U > p_F$. The subsequent stages of the envelope game remain unchanged: in the second stage, player 1 decides whether or not to learn the state of the environment (this decision may now depend on player 1's type); in the third stage, player 2 decides whether to engage in an interaction (depending on whether or not player 1 looked at the state); and in case of an interaction there is a fourth stage in which player 1 decides whether to cooperate. The payoffs of the players may now too depend on player 1's type; they are b_{it} after cooperation and d_{it} after defection, with $i \in \{1, 2\}$ and $t \in \{F, U\}$.

Again, we can analyze this model by characterizing the possible equilibria and by performing evolutionary simulations. The respective results are illustrated in Figure 9.3. There we find four different regimes, three of which correspond to the cases observed in the commitment model:

- An opportunism equilibrium, in which both types of player 1 learn the environmental state, are accepted, and cooperate only if the cost is low
- A no interaction equilibrium, in which player 2 rejects all co-players, irrespective of whether or not they learned the state of the environment
- An ignorance equilibrium, in which both types of player 1 commit themselves by not learning the environmental state and by cooperating by default
- A "partial ignorance equilibrium," in which only a favorable player 1 decides not to learn the state of the environment, is accepted, and cooperates by default; the unfavorable player learns the current environmental state and is rejected by player 2

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

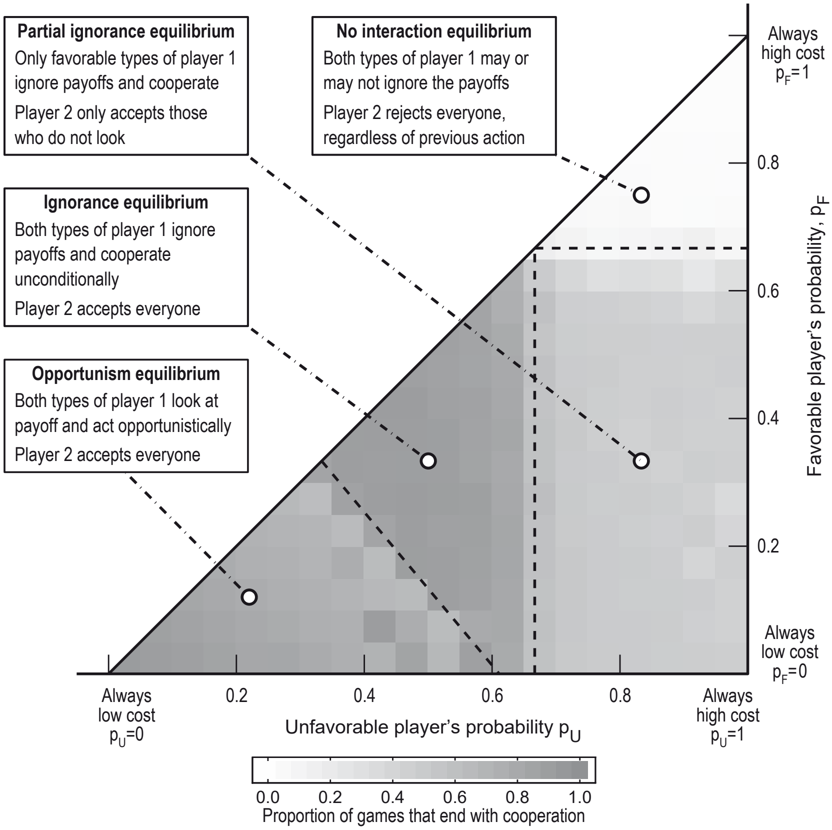

Figure 9.3 Evolution of ignorance in a game with uncertainty about the player types. There are two types of player 1: favorable (F) and unfavorable (U). Favorable players are more likely to encounter a low-cost environment. We consider the evolutionary dynamics that arises in two distinct populations engaged in the envelope game. Population 1 consists of two subpopulations of fixed size, corresponding to the favorable and unfavorable players. Each player 1 knows its own type, but players in population 2 only know the relative abundance of the two types. As before, players are randomly matched, and strategies that yield a higher payoff are more likely to spread within the respective (sub)population (see Appendix 9.1). Depending on the parameter values, we observe that evolution leads to one of four possible equilibria. In the "partial ignorance equilibrium," only favorable players avoid looking into the envelope, whereas unfavorable players look. Player 2 accepts non-looking co-players and rejects all others.

In contrast to the pure commitment model considered previously, the active choice of a player to avoid information can now be crucial. Only if the players themselves have a choice whether or not to ignore their environment can they differentiate themselves from others in the partial ignorance equilibrium. Whereas favorable players can afford to ignore relevant information, as they are likely to end up in a state in which cooperation is mutually beneficial anyway, unfavorable players cannot.

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

As in the baseline model, however, we note that strategic ignorance can only be used as a signal if it is observable. If there were secret ways to learn the true state of the environment, the signal of publicly ignoring information would no longer be costly, and thus no longer reliable. Thus, favorable players can only sustain a partial ignorance equilibrium if they are able to make their ignorance verifiable.

Discussion

Different Dimensions of Strategic Ignorance

The above models illustrate two different mechanisms for how strategic ignorance can emerge, as a way of self-commitment and as a signal. There are, however, further mechanisms. The model of Dubey and Wu (2001), for instance, explores how work performance depends on the intensity of monitoring when a reward is promised to the most productive worker. Workers differ in their baseline productivity. Moreover, their output is subject to random shocks. In such a scenario, employers benefit from showing minimal scrutiny. If, instead, employers collect too much data, workers with a low baseline productivity no longer have an incentive to exert effort. Due to the law of large numbers, their chance to get the reward approaches zero. Kareev and Avrahami (2007) ran a set of experiments which confirms that subjects perform better under minimal, rather than full scrutiny. In other words, employers are better off not finding out as much as they theoretically could.

As another example in which deliberate ignorance occurs, Dana et al. (2007) describe situations in which subjects decide not to learn how their actions affect others. By removing transparency, subjects create a moral wiggle room that allows them to be more selfish. Despite these examples, however, we seem to lack a general theoretical framework that delineates when strategic ignorance pays, and which kind of information is most profitable to ignore.

Deciding Not to Learn versus Deciding Not to Convey

There is an interesting flipside to the problem of deliberate ignorance. Thus far we have considered scenarios in which a focal player decides to remain uncertain about specific aspects of a strategic interaction, although a naive understanding would suggest that resolving the uncertainty should be in the focal player's interest. There are, however, other scenarios in which a focal player deliberately decides to leave others uncertain about certain aspects, even if it seems in the focal player's interest to let them know. As an example, many donors give substantial amounts to charities while deliberately withholding their name. According to the *Chronicle of Philanthropy*, in 2017 there were at least 36 anonymous donations of at least $5 million each in the United States

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

alone. If people donate to gain reputation benefits, why would they decide to leave others ignorant about their good deeds? One possible explanation is that by donating anonymously, donors avoid being harassed by other charities. However, this argument alone does not explain why anonymous donations are often considered more virtuous.

To account for such behavior, Bénabou and Tirole (2006b) propose a signaling model in which players may have three different motives to choose between their actions. The players' decisions may depend on the intrinsic value they attribute to an action, on any extrinsic incentives (such as subsidies), and on the action's reputational value. When players differ in the relative weight they attribute to these three motives, good actions might be suspected of being driven by appearances only. In some cases, players may thus prefer their good actions to be unknown. Similarly, in the signaling model of Hoffman et al. (2018), donors may sometimes have an incentive to deliberately "bury" their positive signals. If such buried signals are eventually revealed, observers not only learn of the donor's good deeds but also that the donor was not interested in public appraisal.

These observations suggest that deliberate ignorance may be just one aspect of a more general class of social quirks, revolving around how we strategically acquire (or ignore) and transmit (or withhold) beneficial information.

Nonstrategic Ignorance

Above, we have given an evolutionary account for why individuals may engage in strategic instances of deliberate ignorance. In the ignorance equilibria, individuals prefer not to know because their ignorance eventually helped them to secure higher material benefits. However, there are also various examples in which individuals avoid information although their ignorance may come at a substantial cost to their long-run welfare, such as when they avoid learning the results of a medical diagnosis (for further examples, see Ellerbrock and Hertwig, Auster and Dana, as well as MacCoun, this volume).

Existing models that account for such behaviors typically assume that subjects have nonstandard preferences (see also Trimmer et al. as well as Brown and Walasek, this volume). Individuals value not only their material payoffs, but also the information they have and how they obtain it. For example, the model by Golman and Loewenstein (2018) can account for many psychologically intuitive behaviors, like natural curiosity and the ostrich effect, by assuming that individual utility not only depends on material payoffs, but also on beliefs and the attention devoted to them.

However, while models based on nonstandard preferences give very reasonable proximate explanations for the psychological mechanisms at work, they typically do not address how subjects have evolved these preferences in the first place. If evolutionary forces have shaped our preferences, it remains unclear why our preferences seem to fail to maximize our material payoffs. To

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

resolve this puzzle, it may be necessary to analyze the preferences we have in light of the ecological context in which they evolved (Fawcett et al. 2014). For instance, a preference to avoid potentially negative information (e.g., disease) could have emerged as a means of self-deception, which in turn may be used to deceive others (Trivers 2011a). In line with this argument, the simulations of Johnson and Fowler (2011) suggest that individuals benefit from a certain degree of overconfidence in evolutionary competitions. Seen from this evolutionary perspective, even nonstrategic instances of deliberate ignorance suddenly have a strategic component.

Appendix 9.1

Static Equilibrium Analysis

For the two models considered in this chapter, one can derive the following equilibrium predictions by backward induction (for the commitment model) or by solving for the Perfect Bayesian Nash equilibria (for the signaling model) (Fudenberg and Tirole 1998).

In the commitment model, the sequential game illustrated in Figure 9.1 allows for three generic outcomes:

1. Opportunism equilibrium: If $p<b_2/(b_2-d_2)$, the game has a unique equilibrium according to which player 1 looks at the environmental state in the second stage, player 2 accepts looking in the third stage, and player 1 cooperates in the fourth stage if and only if the costs are low.
2. Ignorance equilibrium: If $b_2/(b_2-d_2)<p<(b_1-d_1-c_L)/(c_H-c_L)$, there is a unique equilibrium according to which player 1 refuses to look at the environmental state in the second stage, player 2 only accepts those co-players in the third stage who refuse to look, and player 1 unconditionally cooperates in the fourth stage.
3. No interaction equilibrium: If $p>\max\{b_2/(b_2-d_2),(b_1-d_1-c_L)/(c_H-c_L)\}$, then in any equilibrium, player 2 rejects player 1 in the third stage irrespective of player 1's decision in the second stage.

For the signaling model, a full description of possible equilibria is more elaborate. Here, we only describe the equilibria that can be observed for the parameters used in Figure 9.3.

1. Opportunism equilibrium: There is a pooling equilibrium in which both types of player 1 look at the environmental state and are accepted if $p_U\leq 11/15-6/5\,p_F$. This condition ensures that player 2 finds it, on average, beneficial to interact with fully opportunistic co-players.
2. Ignorance equilibrium: There is a pooling equilibrium in which both types of player 1 refuse to look at the state, are accepted, and cooperate unconditionally, if $1/3\leq p_U\leq 2/3$. The first inequality ensures that player 2 has an

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

incentive to reject an interaction with a co-player who looks; the second inequality ensures that both types of player 1 cooperate by default.

3. No interaction equilibrium: If $p_F \geq 2/3$, then in any equilibrium of the game, player 2 rejects her co-player in the third stage.
4. Partial ignorance equilibrium: If $p_F \leq 2/3 \leq p_U$, there is a separating equilibrium in which only the unfavorable players look at the state, only those players who do not look are accepted, and accepted players cooperate unconditionally.

Evolutionary Analysis

The results in Figure 9.2 and Figure 9.3 are based on simulations of a "pairwise comparison process" (Traulsen and Hauert 2009). In the following we describe the process for the signaling model; the process for the commitment model follows by setting $q = 1$ (such that effectively only one type of player 1 is present).

There are two populations: population 1 and population 2 of size N_1 and N_2, respectively. Population 1 consists of two subpopulations: a subpopulation of size qN_1 of unfavorable players and a subpopulation of size $(1-q)N_1$ of favorable players. In each time step, players of population 1 are matched with players in population 2 to play the envelope game. To this end, each player in population 1 is equipped with a strategy represented by a 4-tuple $(x, y_0, y_H, y_L) \in \{0,1\}^4$. The first entry x is the player's probability to learn the state of the environment in stage 2. The other entries give the player's cooperation probability in the fourth stage, given that the costs are unknown (y_0), known to be high (y_H), or known to be low (y_L). Similarly, strategies of players in population 2 are represented by a 2-tuple $(z_l, z_n) \in \{0,1\}^2$. The two entries give the player's probability to accept the co-player, depending on whether or not the co-player looked at the state of the environment. Given each player's strategy in either of the two populations, we can calculate each player's expected payoff.

After interacting in the envelope game, we assume that one player is randomly chosen from either of the two populations. This player is then given the chance to revise his or her strategy. With probability μ (akin to a mutation rate) the player simply picks a new strategy at random. With probability $1 - \mu$, the player compares his or her own payoff with the payoff of a randomly chosen role model from the same (sub)population. If the focal player's payoff is π and the role model's payoff is π', the focal player adopts the role model's strategy with probability $\rho = (1 + \exp[-\beta(\pi' - \pi)])^{-1}$. The parameter $\beta > 0$ is called the strength of selection. In the limiting case $\beta \rightarrow 0$ the imitation probability simplifies to $\rho = 1/2$, independent of the players' payoffs. In that case, imitation occurs essentially at random. For higher values of β, imitation events are increasingly biased in favor of strategies that yield a higher payoff.

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599

This basic process consisting of mutation and imitation is then repeated over many time steps. Figure 9.2a–c shows for each time step of the simulation how often players in population 1 look at the state of the environment on average, and how often the game ends by player 1 cooperating. Figure 9.2d and Figure 9.3 show the respective time averages. The simulations for the commitment model are run with the following parameter values: $b_1=b_2=6$, $d_1=1$, $d_2=-12$, $c_H=7$, $c_L=1$, $N_1=N_2=100$, $\beta=1$. For the signaling model we set $q=0.5$, $b_{1F}=b_{2F}=6$, $b_{1U}=b_{2U}=5$, $d_{1F}=1$, $d_{2F}=-12$, $d_{1U}=-1$, $d_{2U}=-10$, $c_H=7$, $c_L=1$. Other parameter values and alternative evolutionary processes yield similar results.

From "Deliberate Ignorance: Choosing Not to Know," edited by Ralph Hertwig and Christoph Engel.
Strüngmann Forum Reports, vol. 29, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 9780262045599