# 8

# Models of Deliberate Ignorance in Individual Choice

Gordon D. A. Brown and Lukasz Walasek

## Abstract

This chapter reviews models of deliberate ignorance and argues that models developed in both psychology and economics may be useful in understanding different aspects of deliberate ignorance. Such models must specify what quantity is increased at the expense of the potential benefits of the ignored information. A model classification is developed based on the quantity that different models assumed to be so increased. Three broad classes of relevant models are identified: (a) models that assume that utility associated with the content of beliefs may be increased by deliberate ignorance, (b) models that assume that the consistency of beliefs with each other or with a sense of identity may be increased by deliberate ignorance, and (c) models that assume that the quality of decision making may be increased by deliberate ignorance. Gaps in the literature are identified. In particular, it is suggested that insufficient attention has been given to the distinction between the effects on an agent's utility of *acquiring* information (a one-off change) and *possession* of information (being in a steady-state of changed beliefs). Ultimately, models of deliberate ignorance will need to address the relationship between people's (often partial and contradictory) knowledge about the world and their reasoning about that world.

## Introduction

This chapter reviews, with a broad brush, disparate computational and mathematical models which we believe may be useful in understanding deliberate ignorance. We limit our discussion to quantitative and mathematical models rather than descriptive or verbal ones and avoid literature that is primarily empirical. We develop a classification of models and attempt to integrate and evaluate the strengths and weaknesses of different approaches to the modeling of deliberate ignorance.

For our discussion, we adopt a working definition as follows: deliberate ignorance is the conscious individual or collective choice not to seek or use

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

information in situations where the marginal acquisition costs are negligible and the (individual or social) benefits are potentially large. We confine our discussion to decision making at the level of the individual rather than the group, as strategic considerations are the focus of other authors in this volume. Moreover, we do not restrict ourselves to cases where a decision is "deliberate" in the strict sense of "conscious; the result of deliberative thought." Such a restriction would exclude many relevant models in both psychology and economics as well as possibly marginalizing some of the classic examples of deliberate ignorance. It seems far from clear, for example, that avoidance of medical tests always reflects deliberate and conscious processing. More specifically, much research in social and cognitive psychology supposes that we are often unaware of (and may also be mistaken about) the reasons for our actions and inactions; deliberation often seems to follow, rather than precede, a decision. To the extent that this supposition is correct, models of conscious processes alone may miss crucial insights pertinent to deliberate ignorance. Another reason for interpreting "deliberate" loosely is that many of the models most relevant to deliberate ignorance are to be found in the economics literature, and such models are typically interpreted as "as-if" accounts. As-if models do not claim to characterize the deliberative and conscious psychological processes that underpin decision making; instead, they are couched at a higher (e.g., algebraic) level of description and typically aim to make sense of behavior by identifying inferred preferences that explain a person's behavior as being consistent. Thus, there is no assumption of conscious deliberative processing in such models. Here, however, we argue that as-if models nevertheless offer key insights into deliberate ignorance.

The working definition of deliberate ignorance involves the choice not to seek or use information in situations where possession of the information would confer "large potential benefits." However, in seeking to understand why people engage in deliberate ignorance, theorists are necessarily looking for some quantity that is optimized or at least increased as a result of the choice not to look for, or use, additional information (while acknowledging that there may also be many costs). Any model of deliberate ignorance must assume that the ignorance is in some respect beneficial for the agent who seeks it, and the task of the modeler is to identify what the quantity being optimized is, as well as perhaps to specify the relevant psychological mechanisms, either analytically or through simulation. Indeed, the question of "what is being maximized when deliberate ignorance occurs" is central to the classification of models that we develop below. To put it in the language of economic models, people can be understood as having preferences, and the job of models is not to say whether those preferences are right or wrong (for nothing can be said about that; "people want what they want") but simply to identify what those preferences are. The task for an economic model, therefore, is to identify the preferences that deliberate ignorance is helping to satisfy. The approach, therefore, contrasts with evolutionary or functional

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

perspectives, according to which preferences can be explained in terms of contributions to fitness.

Such preferences must act in opposition to preferences *for* information. It is well established that under many circumstances people place a value on gaining information per se, even if the resulting information is unlikely to inform their future actions. Thus, people will pay to discover the secret of a magic trick or to discover what would have happened if they had made different choices in an experiment. People's choices of information in logical reasoning tasks can be better explained if it is assumed that they choose items that will maximize information gain rather than apply logical rules (Oaksford and Chater 1994). In George Miller's terms, we are "informavores" (see also Loewenstein 1994). Demonstrations of deliberate ignorance seem to show that there is some factor deriving from the expected content or consistency of belief, rather than the amount of information gained, which can override this general psychological preference for information seeking. So, what are the preferences that might be satisfied by deliberate ignorance?

## A Preliminary Classification

Figure 8.1 illustrates the sources of deliberate ignorance effects assumed by the various models that we review below. Current beliefs and preferences (i.e., those that hold at the time of decision making) are illustrated on the left side of the figure. Current beliefs may be associated with utility to the extent that they are consistent with an individual's preferences; holding a belief that I carry the gene for Huntington disease may not sit well with my preference for a state of the world in which I have a long and healthy life, and hence changing, suppressing, or simply not thinking about that belief may increase my well-being. The left-hand side also shows the importance for utility of having beliefs that are consistent with one another and/or consistent with one's sense of identity and/or ego; many economic models (reviewed below) accord a central role to identity, consistency, and the possession of positive views about the self.

To identify additional possible sources of deliberate ignorance, Figure 8.1 also represents anticipated future beliefs and preferences (right side of the figure) together with the temporal trajectory linking present and future beliefs. These may also influence decision making at the time it occurs. First, deliberate ignorance may be predicted by a focus on the utility anticipated to be associated with potential future beliefs. As with current beliefs, anticipated possible future beliefs may be associated with the loss of utility to the extent that they are inconsistent with anticipated future preferences, each other, or an anticipated future sense of identity.

Although deliberate ignorance is normally assumed to reflect some anticipated difference between current and future states, only some potential accounts of deliberate ignorance take explicit account of the amount of time for

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

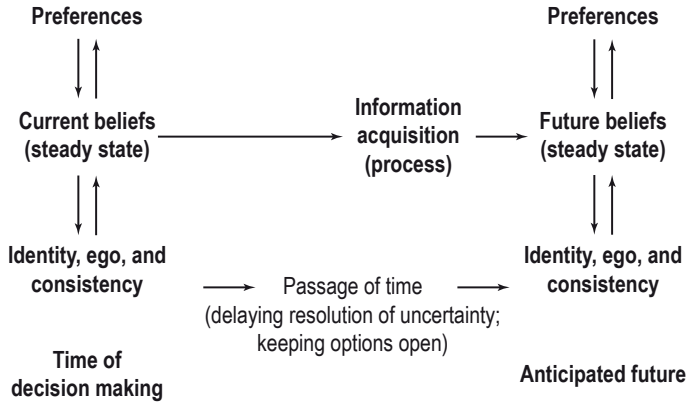


**Figure 8.1** Schematic to illustrate the various sources of deliberate ignorance effects that have been assumed by models.

which a state of ignorance is maintained. Anticipatory emotions, such as hope, may lead to deliberate ignorance if that ignorance increases the amount of time before hope is likely to be dashed (e.g., one might postpone finding out the outcome of a low-probability but high-win gamble, such as a lottery entry). Thus, we view a choice to delay the acquisition of information as a form of deliberate ignorance, because such a choice maintains a state of relative ignorance for longer than necessary. Delay-related deliberate ignorance can also result from models that assume preferences over the timing of uncertainty resolution, or Kreps-Porteus preferences (Kreps and Porteus 1978). We therefore include "passage of time" in Figure 8.1.

In addition, by including "information acquisition" as a separate component, the schematic reflects a distinction between effects on anticipated utility that are due to the *acquisition* of new knowledge and effects due to the *possession* of it. Imagine, for instance, deciding whether to go to the doctor to receive results of a gene test for Huntington disease. It is intuitively plausible that your feelings about going to the doctor would be strongly influenced by a vivid mental picture of yourself sitting in the chair in the doctor's surgery receiving the news, and less influenced by considerations of how you might feel in the longer term as a person in possession of the unfavorable diagnosis, but having had time to adapt to it. Although many models of deliberate ignorance fail to note this distinction, it is important psychologically and is captured in models that assume that necessarily transient emotions, such as surprise, may motivate deliberate ignorance.

Finally, although Figure 8.1 focuses on preferences and beliefs rather than the choices and decisions that would result from them, deliberate ignorance may also be motivated by expectations about the quality of decision making and choice conditional on the amount of information available to the decision maker. Although these expectations are not represented, there are many cases where ecologically

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

optimal decision making or future prediction may be improved by discarding or ignoring information, and we review such models below.

We can now identify three broad categories of models, with the categories differentiated by the quantity that is assumed to be amenable to increase through deliberate ignorance:

1. People may choose to ignore information likely to support beliefs that in some way threaten their preferences (here broadly interpreted to include desires and attitudes). These models assume that utility associated with the *content* of beliefs is maximized.
2. People may ignore information to improve the *consistency* of their beliefs with each other or with their identity; there might be a cognitive cost to inconsistency per se or to believing, for example, that one has low ability if such a belief is inconsistent with the belief that one has high ability. Relatedly, there may be a cognitive cost to changing one's mind and it may be this cost that is minimized by deliberate ignorance.
3. People might be maximizing the *quality of their decision making*. For example, they might base decisions on smaller-than-available samples of information, leading to superior identifications of contingencies in the world albeit at some cost of false positives; ignore information to prevent known cognitive biases contaminating their decision making; and discard information to prevent "overfitting" of predictive models of the environment.

Overall, our classification of computational models is based on their underlying assumptions of how deliberate ignorance may emerge due to interactions involving people's preferences, beliefs, and the time course of resolving uncertainty. This framework is, therefore, distinct from existing taxonomies of deliberate ignorance (Gigerenzer and Garcia-Retamero 2017; Hertwig and Engel 2016). In these other taxonomies, the primary goal is to delineate various causes of deliberate ignorance, but this does not require reference to the mathematical and formal aspects of the underlying processes. How does the model-based taxonomy developed here map onto those developed by others? We see a close correspondence between our "belief-content" models and two of Hertwig and Engel's subcategories of functions of deliberate ignorance (emotion regulation and regret avoidance; suspense and surprise maximization). In addition, models based on enhancing decision-making quality fall neatly within their performance-enhancing subcategory. Our model-derived category of models involving consistency and identity, in contrast, does not fit well although it has some overlap with Hertwig and Engel's "strategic" category as applied to individuals (e.g., self-disciplining).

We now review models within each of these three categories. Perhaps surprisingly, the majority of relevant mathematical models can be found in the economics literature. Indeed, the idea that beliefs can, in themselves, carry implications for current and predicted well-being has received more attention,

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

at least as far as provision of specific models is concerned, in economics than in psychology.

## Deliberate Ignorance and Models of the Content of Beliefs

Several economic models assume that people derive utility from their beliefs about states of the world (Brunnermeier et al. 2017; Brunnermeier and Parker 2005; Caplin and Leahy 2001, 2004; Ely et al. 2015; Epstein 2008; Golman and Loewenstein 2016; Golman et al. 2016, 2017; Köszegi 2003, 2006, 2010; Köszegi and Rabin 2009; Loewenstein 1987). The notion of belief-dependent utility represents a strong departure from the standard approach, according to which beliefs and preferences are independent. In the classic view, a person should only choose to obtain new information for its instrumental value. In utility-from-beliefs models, beliefs about future outcomes or the present state of the world can be a source of positive utility in themselves. Models that explicitly accord a role to belief-related utility necessarily open up the possibility that utility might be increased through deliberate ignorance. Some theorists explicitly explore such implications, while others do not.

Below, we distinguish between content-based models that do and do not allow utility to be influenced by the amount of time that an individual is in a particular belief state ("duration-dependent" and "duration-independent" models). The passage of time could be relevant to deliberate ignorance either because positive emotions (like suspense) might be more valuable if they obtain for longer or because it might be preferable to experience negative emotions, perhaps related to uncertainty, for as short a time as possible or in the future rather than the present.

### Duration-Independent Content-Based Models

A number of economic models quantify the utility loss that may be experienced when preference-relevant beliefs change as a result of new information, and hence can explain deliberate ignorance for such information. Such models typically do not assign a major role to the amount of time that passes between the decision to engage in deliberate ignorance and (potential) information acquisition. For example, Köszegi (2003) invoked utility defined over beliefs to explain why patients may rationally choose to avoid new information about their health condition. According to his model, a patient who learns new information can choose appropriate treatment, which increases anticipatory utility since the patient will expect their health to improve. This increase, however, may be offset by the negative impact of learning that the state of health is poor. The patient must trade off the utility loss associated with receiving bad news against the benefit of a better knowledge of their health, and a decision to avoid visiting the doctor may reflect this trade-off

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

(see also Schweizer and Szech 2018). Köszegi (2010) discusses at length the role of disappointment aversion in informational preferences (which can include deliberate ignorance). Imagine that you possess an instant-lottery ticket with a 50% chance of winning £50 and a 50% chance of winning £100, and the choice of resolving the lottery immediately, or waiting to do so. Under ignorance, you will likely be either surprised or disappointed by the outcome, and the degree of surprise or disappointment will depend on the reference point given by your expectations. The state of ignorance may carry higher or lower utility depending on your degree of disappointment aversion (and possible trade-offs with optimism).

A number of models address the relationship between investment behavior and disappointment aversion. For example, in the Andries and Haddad (2017) model, increases in the subjective probabilities of disappointing outcomes may lead people to prefer infrequent "bundles" of information over more frequent small amounts of information (e.g., checking the performance of investments frequently); for an account based on "news utility," see Pagel (2018). In a series of papers, Golman and Loewenstein have proposed a unified utility-based theoretical framework that captures preferences for acquisition and avoidance of information (see Golman and Loewenstein 2016; Golman et al. 2017, 2019). The key deliberate ignorance-relevant assumption of the model is that people prefer to avoid attending to unpleasant anticipated outcomes. Despite a preference for seeking information for its instrumental value, people may engage in deliberate ignorance if the new information would be psychologically painful or unpleasant. In a number of other models discussed below, preference for information emerges from people's underlying attitudes toward risk, time, and uncertainty. In the Golman and Loewenstein model, however, preference for information is the source of (and can therefore influence) preferences for risk and uncertainty.

Unlike many traditional utility-based models in which utility is assigned to material outcomes, Golman and Loewenstein's model assigns utility to cognitive states, which include (a) strength of attention paid to unanswered questions and (b) subjective judgments about the probabilities of possible answers to such questions. More formally, consider a question set $Q = \{Q_1, \ldots, Q_m\}$ with corresponding attentional weights, $\mathbf{w} = \{w_1, \ldots, w_m\}$. For each question $Q_1$ an individual holds subjective beliefs over potential answers, $\mathcal{A}_i = \left\{A_i^1, A_i^2, \ldots\right\}$. The space of answers and prizes ($X$) is then given by $\alpha = \mathcal{A}_1 \times \mathcal{A}_2 \times \ldots \mathcal{A}_m \times X$. The cognitive state is then given by a subjective probability ($\pi$) defined over possible answers to a given question ($\alpha$) and the vector of attention weights ($\mathbf{w}$). In the model, acquisition of new information is treated as a decision to accept a lottery over possible cognitive states. The utility function in the model is defined over cognitive states that may result from actions $s \in S$, which may lead to a discovery of new information and revision of one's prior beliefs:

$$U\left(\pi, \mathbf{w} | S\right) = \max u\left(s \in S \cdot \left(\pi, \mathbf{w}\right)\right), \tag{8.1}$$

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

where $u$ is the utility associated with a particular action and $U$ is the utility resulting from choosing the action $s$ associated with the maximum $u$. The desire to acquire new information is therefore given by the difference in the expected utility before and after receiving new information:

$$D_i = \sum\nolimits_{A_i \in A_i} \pi_i^0 \left(A_i\right) U\left(\pi^{A_i}, \mathbf{w}^{A_i} | S\right) - U(\pi^0, \mathbf{w}^0 \mid S), \tag{8.2}$$

where $(\pi^0, \mathbf{w}^0)$ specifies the initial cognitive state. Thus, for example, a decision to find out about the results of a medical test depends on the utility associated with the potential cognitive states. If the anticipated result is negative, an individual will avoid the answer to the relevant question $Q_i$. This utility function instantiates three mechanisms that guide the desire to acquire or avoid information:

- Information has instrumental value: its availability can increase the utility of subsequent actions.
- Individuals gain utility from acquiring new information and closing information gaps. A "curiosity motive" is thus conveyed by utility gained from new information absent its instrumental value.
- Motivated attention (and therefore deliberate ignorance) can emerge if the new information influences attention weights.

Individuals may actively avoid information that increases how much they think about unpleasant outcomes, regardless of the instrumental value of new knowledge, and potentially overcome the curiosity motive. Golman and Loewenstein (2016) describe conditions under which motivated cognition can result from the influence of attention weights on utility.

Full specification of the model, its assumptions and consequences, can be found in the original papers (Golman and Loewenstein 2016; Golman et al. 2917, 2019). Here we note that the key determinant of deliberate ignorance in their model is the valence of the surprising information. Surprising information (causing large revision in prior beliefs) is associated with a shift in attention weights upon acquiring new knowledge. This will amplify people's reluctance to gain information when the information is associated with negative valence.

### Duration-Dependent Content-Based Models

A number of other models assign a more central role to the amount of time that passes before uncertainty is resolved. Thus, a preference for deliberate ignorance could result from attempts to maximize positive anticipation or minimize dread, and it seems reasonable to assume that the utility gains of a month of eager anticipation (e.g., of winning the lottery) are greater than the gains from just a week of experiencing the same anticipation. Alternatively, but relatedly, deliberate ignorance could result from a general preference for delaying resolution of uncertainty.

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

Loewenstein (1987) developed a model in which utility could be gained by delaying consumption and presented illustrative data (e.g., people are willing to pay more for a kiss that is delayed by three days than for an immediate one). People may also sacrifice currently preferred consumption options to keep future options open, thus taking account of the fact that their preferences may change in the future (Kreps 1979). Preferences for delaying consumption or commitment (choice) do not in themselves lead to deliberate ignorance. However, models of preferences for delay can be extended to deliberate ignorance when the future outcome is uncertain, because a choice to delay the resolution of uncertainty (i.e., delaying acquisition of the knowledge about an outcome) is a case for preferring ignorance, at least for some period of time. An individual may thus choose to avoid information, such as learning the outcome of a lottery, to maximize the utility derived from suspense or excitement. Chew and Ho (1994) present evidence that hope (which they define as enjoyed maintenance of a state of uncertainty, often regarding a potential gain) is stronger for a low-probability gain.

Formal utility-based models have been developed to capture the role of anticipatory emotions (Caplin and Leahy 2001; see also Dillenberger 2010; Köszegi and Rabin 2009). Of particular relevance to deliberate ignorance are models in economics that build on the axiomatic approach to studying dynamic resolution of uncertainty proposed by Kreps and Porteus (1978). This approach brings preferences for the time at which uncertainty is resolved within the traditional economic utility-based framework. Extensions of this recursive expected utility model have led to multiple belief-based models in economics. For example, the framework has been extended to capture the role of anticipatory anxiety (Caplin and Leahy 2001) as well as the value of suspense and surprise (Ely et al. 2015). Note that the difference between "suspense" and "surprise" mirrors the distinction, raised above, between being in a state of ignorance over time (suspense) and the resolution, at a single time point, of uncertainty (potential surprise, depending on expectations). Ely et al. (2015) use their modeling approach to study the conditions under which noninstrumental information may be sought to maximize surprise and suspense. In their model, an individual may choose a particular information policy to achieve optimal suspense and surprise as beliefs evolve over time. Suspense arises when the uncertainty in an outcome is higher in a future period than in the current period. Surprise is defined as the extent to which beliefs change. This modeling framework captures an important determinant of deliberate ignorance. In many cases people may choose to delay resolving uncertainty to minimize negative and maximize positive emotions. Thus, their model can simultaneously capture cases in which people seek irrelevant (noninstrumental) and avoid relevant (instrumental) information.

Several other models also assume that there is a benefit to optimism and the anticipation of positive future events, such as passing an examination or becoming rich due to the success of one's investments. In Brunnermeier and

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

Parker (2005), individuals are assumed to hold incorrect and yet optimal (in the sense of being happiness-maximizing) overoptimistic beliefs. Over a long enough time period, the negative cost (for decision making) of being too optimistic may be outweighed by the positive utility of holding an erroneous belief. A decision not to seek information can therefore be motivated by the individual's desire to maintain a positive outlook on the future (cf. Köszegi 2010). The idea that "living with risk" can be associated with anxiety or hope is also explored in a model developed by Epstein (2008), who shows that such a model can predict information avoidance when an unfavorable outcome is very likely together with information seeking when a favorable outcome is anticipated. A related account was developed by Bénabou (2013), who describes a model of groupthink and shows that denial of negative information can either be contagious or self-limiting, depending on how harmful to others it is. More specifically, if the other members of an agent's group engage in deliberate ignorance (of bad news), they may act in a way that is either good for the agent (thus reducing the agent's own incentive to increase anticipatory utility by engaging in deliberate ignorance) or bad for the agent (increasing the agent's own incentive for deliberate ignorance). Individual incentives for deliberate ignorance, therefore, depend both on the accuracy of other's beliefs and on the probabilities of good and bad outcomes to collective action.

Anticipated regret may be a key affective consideration that underpins choosing not to know. Gigerenzer and Garcia-Retamero (2017) proposed a modeling framework to explain why people may choose to avoid information about both positive and negative outcomes. In their regret theory of deliberate ignorance, individuals are assumed to avoid the maximum possible anticipated regret (minimax regret criterion). In the model, the emotions associated with the possible outcomes of acquiring knowledge may encourage a person to prefer ignorance. For example, a person may choose not to know when they will die if one of the possible answers is that they will die very soon, thus causing high anticipated regret. The model can be adapted to account for positive emotions, as when the anticipated regret is based on the loss of suspense and surprise. An individual may choose to ignore new information to maintain these positive emotions.

## Deliberate Ignorance and Models of the Consistency of Beliefs and Identity Maintenance

In the previous section, we discussed how deliberate ignorance can be understood in terms of maximizing positive emotions (such as anticipation of favorable outcomes) and minimizing inconsistency between individuals' beliefs and their preferences and desires. However, deliberate ignorance may emerge from a preference for belief consistency, irrespective of the content of those beliefs. Closely related models assume that utility is gained by maintaining a

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

consistent or a positive *identity*. Here, we briefly review this class of model and show how consistency preference and ego protection models may shed light on deliberate ignorance. We note that a preference for consistency is different from assuming loss aversion in preferences over *changes* in beliefs, as is assumed in the model of Köszegi and Rabin (2009).

The idea that consistency matters has a long history in social and clinical psychology. For example, Heider's Balance Theory maintains that unbalanced structures of cognitions produce negative affect (Heider 1958). Cognitive dissonance is assumed to arise when attitudes, social norms, and behavior are not aligned (Festinger 1954), and an influential line of research argues that, rather than having direct access to our preferences, we infer them from our own behavior (Bem 1967; Wilson 2002). Effects of cognitive dissonance have also been noted in the economics literature (Mullainathan and Washington 2009), with Golman et al. (2016) providing a comprehensive review of the evidence for preferences for "belief consonance" across, rather than within, individuals.

A concern for consistency could motivate deliberate ignorance in cases where the ignored information might threaten an existing worldview. The phenomenon of confirmation bias can be seen as a form of deliberate ignorance. Confirmation bias occurs when people pay selective attention to evidence that is consistent with their existing attitudes or beliefs at the expense of ignoring information inconsistent with those attitudes and beliefs. Confirmation bias can thus be seen as an intermediate point on a continuum with complete attention at one extreme and deliberate ignorance at the other. Although there are few formal models within the social and clinical literatures that have given rise to concepts such as cognitive dissonance, some relevant models exist in both economics and cognitive psychology. Many of these models have not been directly applied to deliberate ignorance, so we review them briefly as they provide a framework within which deliberate ignorance could be accommodated.

Falk and Zimmerman (2017) describe a model in which individuals have a preference to behave consistently, and hence for having consistent beliefs (on the assumption that consistent beliefs are more likely to lead to consistent behavior). The preference is assumed to reflect the adaptive value of signaling strength to others and is captured in a utility term that represents observers' certainty about the relevant agent's beliefs. Deliberate ignorance offers one, although not the only, way to maintain consistency of one's beliefs. It may, in turn, be easier to behave consistently if one's beliefs are consistent. This approach adds to a number of economic models of cognitive dissonance (e.g., Konow 2000; Rabin 1994) to suggest that people can, at least to some extent, control their beliefs. They are consistent with the idea that one way in which dissonance may be reduced is through deliberate ignorance (e.g., Akerlof and Dickens 1982).

Yarif (2002) believes an individual's utility function is assumed to capture a trade-off between (a) utility gained from making good decisions, as in the standard model, and (b) utility gained from having consistent beliefs. Thus,

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

agents can "choose what beliefs to hold." Yarif shows that agents may prefer to avoid new information if the cost (to consistency) is greater than the benefit (to improved decision making).

Bénabou and Tirole (2011) present one of the many related economic models which accord an important role to *identity*. In such models, "identity" typically refers to a person's self-image as well as their feelings about themselves, and utility can be gained or lost by making choices that are or are not identity consistent. In the Bénabou and Tirole model, people infer their values from their past choices, and invest in and protect their identity. One way of doing so is by avoiding markets (or even thoughts about markets) that involve prices being placed on goods (e.g., sex, votes, bodily organs) which, according to the given self-identity, should not be bought and sold. These types of models thus suggest that deliberate ignorance may serve a role in value preservation and/or avoiding feelings such as guilt, disgust, and repugnance (Roth 2007).

Along related lines, deliberate ignorance may help people preserve their sense of moral worth. A large body of findings suggests that deliberate ignorance of the effects of one's actions on others can allow one to maintain one's self-image without sacrificing one's own payoffs. This relates to the idea of "moral wiggle room." Grossman and Van Der Weel (2017b) describe a model of the interplay between altruistic preferences, selfish tendencies, and selfish preferences. They use the model to determine the conditions under which there can be an "ignorance equilibrium," whereby ignorance limits availability of information to a person that can be detrimental to their self-image (see also Serra-Garcia and Szech 2018).

According to some models, deliberate ignorance may also arise as a result of a preference for believing that one has high ability, or "ego utility" (Köszegi 2006). Köszegi's model shows how a decision maker who is happy with the ego-related beliefs that they currently hold has an incentive to avoid receiving information that might threaten those beliefs: the "self-image protection motive" (see also Johnson and Fowler 2011). Other identity-related models show how deliberate ignorance may be used as a device for self-control. For example, Carrillo (2005) shows that an individual with time-inconsistent preferences may under some circumstances choose not to obtain information about the future consequences of actions (e.g., the adverse health consequences of smoking, or the pleasure derived from a certain type of consumption) because they fear that new information might cause them to behave in a less healthy way in the future (see also Carrillo and Mariotti 2000). For example, learning that cocaine consumption is enjoyable might lead to a present preference for immediate consumption but future abstinence. However, due to temporal discounting, longer-term overconsumption might be anticipated and hence avoiding knowledge about the present value of cocaine consumption may be beneficial.

Another approach to understanding deliberate ignorance can be found in parallel constraint satisfaction (PCS) modeling. PCS models are a class of connectionist models in which cognitive processes are represented by spreading

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

activation in a network of interconnected nodes. In the context of motivated cognition, nodes are taken to represent goals, actions, or beliefs, and the connections between these nodes reflect the strength of association or level of compatibility. PCS models involve feedback relations which allow for satisfaction of multiple simultaneous constraints imposed on a network (Read et al. 1997). The PCS process attempts to find the highest level of organization (or lowest energy) in activation of the nodes, given specific relations between the nodes (imposed by the researcher). The level of organization or harmony in the network has been interpreted as a measure of cognitive consistency (e.g. coherence of beliefs).

How could a PCS approach shed light on deliberate ignorance? Thagard (1989) describes a model in which the relation between propositions and observations contributes to the overall coherence of an entire system of beliefs (see also Thagard 2006). Using a neural network implementation, any two propositions can be mutually excitatory if they are coherent with each other (if *A* and *B* cohere, then *B* and *A* cohere) or mutually inhibitory if they are not coherent (if *A* contradicts *B*, then *A* and *B* do not cohere). Simulating multiple propositions and observations, the model has been used to explain phenomena such as the acceptability of Copernicus's theory of the solar system (Nowak and Thagard 1992). With an extension of the model, it is possible to account for seemingly irrational beliefs. For example, the emotional value of the links between propositions can lead to a situation in which individuals engage in self-deception (Sahdra and Thagard 2003). In another extension, coherent propositions about greenhouse gases and their role in global climate change can be rejected if they conflict with one's values, such as the importance of a small government. Thus, a coherence model provides a natural perspective on cases of deliberate ignorance, such as when an individual actively chooses to avoid information that could disturb the coherence of their existing beliefs. If we assume that coherence and self-deception are important determinants of subjective well-being (Sahdra and Thagard 2003), then it may be psychologically beneficial to avoid certain sources of information to avoid the risk of challenging the existing belief system.

A final type of model that we consider under this discussion of consistency and identity focuses on the tension that may arise when people's attitudes and attitude-related beliefs are in conflict with a social norm. Under such conditions, people may choose to ignore sources of evidence (e.g., newspaper articles or people with uncongenial opinions) to reduce the tension between expressing attitudes that are consistent with their own beliefs and expressing attitudes that are consistent with the social norm. In social sampling theory (Brown et al. 2019), individuals are assumed to have their own authentic private attitudes (e.g., political attitudes) which are not visible to others. Individuals are also assumed to be sensitive to the distribution of attitudes held by a social group. The attitude people decide to express publicly represents a trade-off between two opposing forces: an *authenticity preference* which motivates an individual 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

act in a way that is in line with their underlying beliefs and attitudes, and *social extremeness aversion* which discourages them from endorsing attitudes that are too different from those of others in a social context. Since both forces are a source of utility loss, an individual will express utility-maximizing attitudes that will reflect a compromise between personal and social beliefs or attitudes. In such a model, utility-maximizing agents have an incentive to seek information that is consistent with their privately held attitude and to avoid information that challenges it (e.g., by preferring the company of similarly minded individuals). In this way they can better satisfy their authenticity preference without being socially extreme.

## Maximizing Quality of Decision Making

The working definition of deliberate ignorance that we have used refers to not seeking or using information when it might be beneficial to do so. As noted in the introduction, however, what counts as "beneficial" is relative to a combination of an agent's goals and the environment in which they live. In particular, the costs and benefits of a strategy may motivate self-perception bias, which in turn can motivate deliberate ignorance. For example, Johnson and Fowler (2011) describe a model in which it is advantageous for an individual to overestimate their own ability. This can occur when (a) an agent's decision whether or not to enter a contest for resources is made on the basis of the agent's beliefs about their own ability, relative to the ability of the other potential combatant, and (b) the positive payoff from winning a contest is much greater (in absolute terms) than the negative payoff that would result from losing the contest. The suggestion that an asymmetrical payoff matrix may motivate self-delusion is distinct from the idea that particular patterns of behavior, such as consistency, may serve a useful role in signaling abilities to others (Falk and Zimmermann 2017). The point here is that payoff contingencies in the environment may lead an adaptively rational agent to perceive the world in a non-veridical way. One way of achieving this is through a form of deliberate ignorance that involves making estimates of some quantity or association using fewer data than are available, even if there is negligible cost to obtaining the additional data. "Quality of decision making" could be quantified in terms of minimizing processing cost or decision time, minimizing false positive errors, or minimizing false negative errors. Several studies from the judgment and decision-making literature focus on avoiding false negatives, to which we now turn.

One example concerns the use of small samples in estimating payouts, which could reflect a deliberate strategy. It is well established in the psychology of judgment and decision making that many judgments of social and economic quantities are based on small samples retrieved from one's memory or the immediate environment (e.g., Fiedler and Juslin 2006). The "small samples" assumption has been used to account for a variety of phenomena,

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

including stereotype formation, illusory correlations, confirmation bias, polarization, and overconfidence. The use of small samples may reflect the increasing cost of expanding sample size, combined with diminishing returns for accuracy of judgment. However, basing estimates on small samples may be adaptive even in the absence of such costs, and hence be an instance of deliberate ignorance. For example, statistical modeling has shown that small samples may be better for detecting small associations in the environment, albeit at the expense of false positives or less accurate assessment of the strength of the association than would be obtained with a larger sample (e.g., Fiedler and Kareev 2006). Thinking in ecological terms, it is easy to envisage circumstances in which it is more important to become aware, at the earliest possible stage, of possible contingences in the world than it is to avoid developing beliefs in relationships that do not in fact exist. Deliberate ignorance (in the form of choosing smaller samples than available) could reflect this need, although the fact that small samples may be better for some purposes is not in itself an instance of deliberate ignorance. (We note that, as mentioned in the introduction, we are including cases where ignorance may reflect adaptive considerations even in the absence of conscious deliberation of benefits of information.)

The "decisions from experience" paradigm provides another case where simulations have shown that deliberate ignorance might help. People seem to choose between payoff distributions on the basis of small (ca. seven) samples from each, even when the cost of obtaining larger samples is relatively small. Hertwig and Pleskac (2010) show that small samples help in choosing between payoffs, because small samples amplify the difference between the earnings associated with uncertain payoff distributions. Their model points to another case where ignorance (in this case of a wider sample) could make choices easier by making options more distinct, even if additional information could be acquired at little or no cost.

Other models of decision making that illustrate how deliberate ignorance can lead to improved performance can be found in *Simple Heuristics That Make Us Smart* by Gigerenzer et al. (2000). Researchers working within this tradition have found that decision rules are, under many circumstances, more successful when they ignore part of the information. In particular, strategies such as tallying (i.e., counting up the number of cues that favor an option, without weighting them) may work well even though part of the information (i.e., cue weighting) is being ignored.

The use of heuristics such as tallying may seem far removed from traditional cases of deliberate ignorance, as they do not seem to involve deliberative and active decisions to ignore certain information. However, as we noted in our introductory remarks, there is no clear dividing line to be made between models which do and do not assume conscious deliberative processing in ignoring information. Therefore, we include them here as an example of how discarding available information can improve judgment and decision making.

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 an example of how forgetting may aid heuristic inference, see Trimmer et al. (this volume).

## Conclusion

In this chapter, we have reviewed a number of different approaches to modeling deliberate ignorance from both economics and psychology. The models reviewed identify a number of different mechanisms through which deliberate ignorance may occur: deliberate ignorance may reflect a desire to maintain beliefs that are consistent with preferences, a desire to maintain a consistent identity or pattern of behavior, a desire to maintain a positive self-image, or a desire to maximize the quality of judgment and decision making relative to a particular set of goals and within an environment of a given structure.

Because the definition of deliberate ignorance restricts it to cases where the cost of acquiring information is negligible, we have omitted from this discussion a broad class of models, mainly developed within economics, on "rational inattention" (e.g., Sims 2003). These models typically assume that there is a cost to acquiring information, and hence account for cases where decision-relevant information is ignored by an optimal decision maker because the costs of acquiring the information exceeds the expected benefits of having it.

What, if any, general conclusions can be drawn? It is evident that many models may shed light on the phenomenon of deliberate ignorance. Indeed, the various models that we have described above cover, between them, most of the types of deliberate ignorance included in the taxonomy developed by Hertwig and Engel (this volume, 2016). One of the points that they make is that psychology needs to pay more attention to deliberate ignorance, and in this context, it is noteworthy that most of the models we have identified are to be found within the economic, rather than the psychological, literature. There is clearly a need for psychologists interested in deliberate ignorance to pay more attention to a number of these papers in the economics tradition. Indeed, a casual examination of the articles that cite economic models of evidently psychological phenomena (e.g., cognitive dissonance, identity, optimism, confirmation bias) reveals a striking absence of articles within mainstream psychology journals. We also note that existing psychology-based taxonomies tend to place more emphasis on fairness motivations than has been seen in models, while the reverse is the case for considerations surrounding identity maintenance. More generally, there is something of a mismatch between emphases in the modeling and non-modeling literature: the former devotes relatively more attention to maintenance of identity and consistency whereas the latter attends more to issues of ensuring fairness.

An open question is the extent to which a unified account of deliberate ignorance can or should be sought. We have tried to show that there are some common features underlying the classes of models. Some emphasize utility

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

that relates to the content of beliefs, whereas others place the emphasis on consistency of beliefs with either each other or with a desired self-image. We have treated these models (those concerned with identity and those concerned with belief consistency) as a single category because beliefs about one's own identity can be thought of simply as another set of beliefs that must be consistent with other beliefs. We note, however, that beliefs about self and identity seem likely to be ones where the content is particularly important, and hence models that refer to utility gained from the content of beliefs may be linked to models that focus on consistency of beliefs per se.

Still other models focus on the length of time over which emotions, such as fear or optimism, are present. At the present stage of theoretical understanding, these appear to be very different and individually plausible sources of deliberate ignorance, and hence the existence of a variety of models adds support to the idea that there is no single determinant of deliberate ignorance (or at least no determinant that is not so general as to be vacuous). Moreover, there may be domain specificity. It is an open question as to whether work on choice under uncertainty in the laboratory will generalize to feelings about health states. However, there is clearly much scope for further competition between and unification of models within each of the classes that we have described, and for the models to be brought into contact with a rich variety of the empirical and theoretical psychological literature.

There is an intuitive distinction between the psychology involved in *acquiring* information and the psychology of being *in the state of having* that information (with the latter bringing the need to consider the time course of adaptation). The acquisition–consumption distinction has already been emphasized in the literature of consumer choice (Hsee et al. 2009); we suggest that models will need to pay more attention to this distinction as it relates to deliberate ignorance than has hitherto been the case. We also note a potential link with the exploration–exploitation dilemma; we can imagine that an organism might be deliberately ignorant to remain in an exploitation state.

Finally, one further avenue for future research is the importance of prior beliefs. In many cases, people must have some idea about the likely valence of information before deciding whether or not to reduce uncertainty by accessing the relevant information. Prior expectations are incorporated in some, but by no means all, of the models we have discussed above.

## Acknowledgments

This study was supported by the Economic and Social Research Council (U.K.) [grant number ES/P008976/1], the Leverhulme Trust [grant number RP2012-V-022], and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program [grant agreement No 788826]. We thank numerous members of the Forum, in particular the editors, for their detailed comments on earlier versions of this manuscript.

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

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