An Information-Gap Framework for Capturing Preferences About Uncertainty

Russell Golman
Department of Social and Decision Sciences
Carnegie Mellon University
rgolman@andrew.cmu.edu

George Loewenstein
Department of Social and Decision Sciences
Carnegie Mellon University
gl20@andrew.cmu.edu

ABSTRACT

We propose an integrated theoretical framework that captures preferences for acquiring or avoiding information as well as preferences for exposure to uncertainty (i.e., risk or ambiguity) by allowing utility to depend not just on material payoffs but also on beliefs and the attention devoted to them. We use this framework to introduce the concept of an information gap – a specific uncertainty that one recognizes and is aware of. We characterize a specific utility function that describes feelings about information gaps. We suggest that feelings about information gaps are the source of curiosity as well as a second motive to manage one's thoughts through information acquisition or avoidance. In addition, we suggest that feelings about information gaps also contribute to risk- and ambiguity preferences.

Keywords

Ambiguity, curiosity, information gap, motivated attention, ostrich effect, risk

1. INTRODUCTION

In a seminal paper titled "The Mind as a Consuming Organ," Thomas Schelling (1987) pointed out that much consumption is not of the material sort, but takes place largely "in the mind." Research in psychology, decision theory, and economics has identified a number of motives underlying informational consumption, from the powerful force of curiosity (Loewenstein, 1994) to the pleasures of knowledge and insight (Karlsson et al., 2004). Moreover, even when missing information is not available to an individual, demand for this information plays a role in decision making under uncertainty. Here we propose a unified theoretical framework that allows us to model feelings about information and about *information gaps* – specific uncertainties that an individual recognizes and is aware of. We present a specific utility function that takes as input beliefs and the attention devoted to them (as well as material payoffs). This utility model can be applied to decision making about information acquisition or avoidance as well as to decision making under risk and ambiguity (as described in Table 1).

Decision about:	Domain of:
Whether to address an uncertainty	Information acquisition or avoidance
Whether to expose oneself to an uncertainty	Risky or ambiguous choice

Table 1: Two domains of decision making affected by feelings about uncertainty.

In one branch of the economics literature, preferences about information have been viewed as derivative from risk preferences (e.g., Kreps and Porteus, 1978; Wakker, 1988; Grant et al., 1998; Dillenberger, 2010). We take a complementary perspective, considering preferences about information as primitive and viewing preferences about risk and ambiguity as derivative of them.

The standard account of preferences about information holds that information is valuable because, and only to the extent that, it enables people to make superior decisions that raise their expected utility (Hirshleifer and Riley, 1979). Often, however, individuals seek information purely to satisfy curiosity, which refers to the desire for information for its own sake – i.e., specifically *not* for its ability to improve decision making. Curiosity correlates with brain activity in regions thought to relate to anticipated reward (Kang et al., 2009), suggesting that information is a reward in and of itself. Loewenstein (1994) proposed an information-gap account of curiosity, and our framework allows us to capture this motive for information acquisition within an expanded utility model. While curiosity is a powerful motive for information acquisition, there nevertheless are many situations in which people actively choose to avoid information, e.g., not obtaining a costless medical test. We hypothesize that information avoidance derives from a desire to avoid increasing attention on a negative anticipated outcome. More generally, we suggest that individuals have an inclination to seek (or avoid) information whenever they anticipate that what they discover will be pleasurable (or painful). Of course, ex-ante beliefs about such events are already good or bad respectively (Eliaz and Spiegler, 2006), but there can be a big difference between discovering something for sure and simply considering it a likely possibility. Our additional assumption is that obtaining news tends to increase attention to it (as in Gabaix et al., 2006; Tasoff and Madarász, 2009), which leads to the implication that people will seek information about questions they like thinking about and will avoid information about questions they do not like thinking about. We explore the implications of our proposed utility model for information acquisition or avoidance in a companion paper (Golman and Loewenstein, 2015a), and we outline this analysis in Section 5.

The standard account of preferences about risk and ambiguity considers these preferences to be primitives in a model (e.g., Anscombe and Aumann, 1963; Klibanoff et al., 2005). However, research has shown that missing information has a profound impact on decision making under risk and ambiguity. For example, Ritov and Baron (1990) studied hypothetical decisions concerning whether to vaccinate a child, when the vaccine reduces the risk of the child dying from a disease but might itself be harmful. When the uncertainty was caused by salient missing information about the risks from vaccination – a child had a high risk of being harmed by the vaccine or no risk at all but it was impossible to find out which – subjects were more reluctant to vaccinate than in a situation in which all children faced a similar risk and there was no salient missing information. In a second companion paper (Golman and Loewenstein, 2015b) we argue that the information-gap concept developed here underlies an alternative account of risk and ambiguity aversion (and seeking) that is conceptually different from, and has different testable implications from, the usual account of risk aversion involving loss aversion and the usual account of ambiguity aversion involving vague probabilities.[1] In Section 6 we outline our argument that salient information gaps can either increase or decrease preference for uncertain gambles depending on whether it is painful or pleasurable to think about the information one is missing.

Our expanded utility model builds on the insights of Caplin and Leahy (2001).[2] Caplin and Leahy recognize that anticipatory feelings about prizes that might be received in the future can affect utility. We follow them (and Kőszegi (2010) as well) in applying expected utility theory to psychological states rather than to physical prizes, but we expand the domain of psychological states that people can have feelings about. In doing so, we incorporate Tasoff and Madarász's (2009) insight that information stimulates attention and thus complements anticipatory feelings. Kreps and Porteus (1978) present a model capturing preferences for early or late resolution of uncertainty, and Dillenberger (2010) captures preferences for one-shot or sequential resolution of uncertainty; this line of research thus deals with when, but not whether, an individual prefers to acquire information. Our model focuses just on the latter issue, but with it one could address the timing of uncertainty resolution by making additional assumptions about time preference.

We rely on a reduced form model of knowledge and awareness to describe information gaps – and the desire to fill them or ignore them – in order to avoid the complications of working with information partitions in a state-space model of knowledge (as in Aumann, 1976). The standard partitional state-space framework permits a distinction between two states of affairs – knowing and not knowing – but makes it difficult to capture unawareness (Modica and Rustichini, 1994; Dekel et al., 1998). We introduce a question-answer knowledge structure that allows us easily to draw an important distinction between *three* different states: knowing (represented by a question and a particular answer); not knowing, but knowing that one doesn't know (represented by a question and a set of possible answers); and not knowing and not knowing what one doesn't know (represented by the absence of an activated question). This third state corresponds to pure unawareness (Li, 2008), in the sense that an individual is unaware of the question itself and does not even distinguish different possible answers. (In contrast, our question-answer structure does not capture partial unawareness, in the sense of an individual being aware of a question and proper subset of possible answers, but unaware of some other remaining possible answers.) The question-answer structure is consistent with, and could be cast in terms of, a generalized state-space model (e.g., Modica and Rustichini, 1999; Heifetz et al., 2006), but we find the question-answer structure convenient to use.

The question-answer knowledge structure is intended to reflect human information-processing capabilities. Our cognitive maps of the world are not sets of possible states, each described in exquisite detail to account for all possible consequences of all possible decisions. Instead, people attend to a few relevant aspects of a situation and use limited information to make a broad judgment that can be refined later, if necessary. People tend to set goals and monitor their progress toward them in order to navigate a complex world (Miller et al., 1960; Locke and Latham, 1990; Loewenstein, 1999). We advance the idea that the acquisition of knowledge is also goal-oriented. We don't simply seek out information to maximize the data available to us or even to optimize future decisions, but instead tend to seek answers to questions that are either posed to us or that we pose to ourselves. Questions are, therefore, very much like informational goals or reference points. Indeed, focusing on a question that one cannot answer – e.g., a puzzle one cannot figure out – can torment a person and at the same time motivate the search for an answer, much as a high reference point can simultaneously detract from utility and motivate one to strive to reach it.

2. THEORETICAL FRAMEWORK

2.1 Cognitive States

Traditional economic theory assumes that utility is a function of consumption bundles or material outcomes, or (perhaps subjective) distributions thereof. Our basic premise is that utility depends not only on such material outcomes but also on one's cognitive state, encompassing the attention paid to each of the issues or questions that one is aware of as well as subjective judgments about the possible answers to these questions. While people have preferences about their beliefs (and the attention paid to them), we do not treat beliefs (or attention) as choice variables. People can choose whether or not to acquire information that will influence beliefs, but we assume that one's beliefs, given one's information, are constrained by Bayesian inference.

While there surely is an infinite set of possible states of the world, we assume, realistically we believe, that a person can only conceive of a finite number of questions at any one time. We represent awareness with an array of '*activated*' questions and a remaining set of '*latent*' questions. Activated ques-

[1]For example, we show that low-stakes risk aversion (Rabin, 2000) could be attributed to the discomfort of thinking about uncertainties.

[2]Many have considered the notion that people derive utility from their beliefs (Abelson, 1986; Geanakoplos et al., 1989; Asch et al., 1990; Yariv, 2001; Kadane et al., 2008).

tions are those that the individual is aware of. Latent questions are those that the individual could become, but is not currently, aware of. The finite subset of questions a person is aware of (i.e., paying at least some attention to) is denoted $\mathcal{Q}$. We label these activated questions as $Q_1, \ldots, Q_m$. A vector of attention weights $\mathbf{w} = (w_1, \ldots, w_m) \in \mathbb{R}_+^m$ indicates how much attention each activated question gets.[3] These attention weights depend on three factors that we designate "*importance*," "*salience*," and "*surprise*." We return to define and discuss these concepts in Section 3.

A question Q_i has a countable set[4] of possible (mutually exclusive) answers $\mathcal{A}_i = \{A_i^1, A_i^2, \ldots\}$.[5] A person may not know the correct answer to a given question, but reasonably has a subjective belief about the probability that each answer is correct.[6] (The subjective probabilities across different questions may well be mutually dependent.) This framework allows us to capture information gaps, which are represented as activated questions lacking known correct answers, as depicted in Table 2.

Question	Answer	Belief	
Latent	–	Unawareness	
Activated	Unknown	Uncertainty	$\updownarrow$ information gap
	Known	Certainty	

Table 2: The question-answer knowledge structure.

Anticipated material outcomes, or prizes, can also be incorporated into this framework. We let X denote a countable set of prizes – i.e., material outcomes. The subjective probability over these prizes is in general mutually dependent with the subjective probability over answers to activated questions; that is, the receipt of new information often leads to revised beliefs about the likelihood of answers to many different questions as well as about the likelihood of different material outcomes. Denote the space of answer sets together with prizes as $\alpha = \mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_m \times X$. Thus, given a state of awareness defined by the set of activated questions $\mathcal{Q}$,[7] we represent a person's cognitive state C with a subjective probability measure π defined over α (i.e., over possible answers to activated questions as well as eventual prizes) and a vector of attention weights $\mathbf{w}$. We denote the set of all possible cognitive states as $\mathcal{C} = \Delta(\alpha) \times \mathbb{R}_+^m$ (with the notation $\Delta(\alpha)$ referring to the space of probability distributions over α with finite entropy. The restriction to distributions with finite entropy serves a technical purpose, but it should not trouble us – intuitively, it means that a person cannot be aware of an infinite amount of information, which is also the basis for our assumption that the set of activated questions is finite.). Each marginal distribution π_i specifies the subjective probability of possible answers to question Q_i, and similarly π_X specifies the subjective probability over prizes.[8]

The formal representation of a cognitive state is depicted in Table 3. Consider, for example, a college professor deciding whether or not to look at her teaching ratings. The set of activated questions (and possible answers) might include: "How many of my students liked my teaching?" $(0, 1, 2, \ldots)$; "Did they applaud on the last day of class?" (yes/no); "How good a teacher am I?" (great, good, so-so, bad, awful); "Will I get tenure?" (yes/no). Prior belief about the first question might be quite uncertain. The answer to the second question, on the other hand, might already be known with certainty. There may or may not be much uncertainty about the third and fourth questions. All of these beliefs (to the extent they are uncertain) are jointly dependent. The material outcome might be next year's salary, which would also depend on (but not be completely determined by) whether or not she gets tenure. Looking at the ratings will definitively answer the first question and may resolve some, but not all, of the uncertainty surrounding the other issues.

2.2 Actions

A decision maker has the possibility of taking actions with two kinds of effects: *informational* actions contribute to subjective judgments about the world by answering a question; and *instrumental* actions affect the chances of receiving various prizes (outcomes). For example, wagering on the color of a ball drawn from an urn is an instrumental action. Examining the contents of the urn is an informational action. Informational actions affect the subjective probability measure through the conditioning of beliefs on the discovered answer. Instrumental actions affect beliefs directly by changing the distribution over prizes conditional on subjective judgments. Both instrumental and informational actions also impact attention weights through their respective effects on importance and surprise. Note that some actions will have both instrumental and informational effects. Examples include paying a fee for a property value appraisal or hiring a private eye.

At any point in time an individual can be characterized by a prior cognitive state consisting of subjective probability measure π^0 and attention weight vector $\mathbf{w}^0$. Actions, in general, are operators on cognitive states that map to new cognitive states or to distributions over cognitive states. A purely instrumental action acting on the prior cognitive state determines a particular new cognitive state. Typically, it preserves the prior subjective judgment about the probability of each answer set and then specifies a new distribution over prizes conditional on each possible answer set. An instrumental action may also affect the importance of various questions (as formalized in the next section) and thereby influence the attention weights. For example, the decision to participate in a karaoke session will likely raise the attention weight on the question "Am I a good singer?"

Acquiring information also changes one's cognitive state. Ex ante, as one does not know which answer will be discovered, the prospect of acquiring information offers the decision maker a lottery over cognitive states. Upon learning answer A_i to question Q_i, one's subjective probability

[3]We can think of the (presumably infinite) set of latent questions as having attention weights of zero.

[4]We use the term countable here to mean *at most countable*. The restriction of a countable set of answers to a countable set of possible questions does still allow an uncountable set of possible states of the world, but as awareness is finite, the precise state of the world would be unknowable.

[5]We assume that there is no such thing as an answer that is disconnected from a question.

[6]By *subjective* probability, we mean personal probability, but we take it to be observable by direct elicitation.

[7]In most cases, we will assume that activation of questions is determined exogenously – i.e., by the environment. We don't model growing awareness (see Karni and Vierø, 2013).

[8]For any $\tilde{\mathcal{A}} \subseteq \mathcal{A}_i$, we have $\pi_i(\tilde{\mathcal{A}}) = \pi(\mathcal{A}_1 \times \cdots \times \mathcal{A}_{i-1} \times \tilde{\mathcal{A}} \times \mathcal{A}_{i+1} \times \cdots \times \mathcal{A}_m \times X)$.

Activated Questions	Possible Answers	Subjective Probabilities*	Attention Weights
Q_1	$\mathcal{A}_1 = \{A_1^1, A_1^2, \ldots\}$	$[\pi_1(A_1^1), \pi_1(A_1^2), \ldots]$	w_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$
Q_m	$\mathcal{A}_m = \{A_m^1, A_m^2, \ldots\}$	$[\pi_m(A_m^1), \pi_m(A_m^2), \ldots]$	w_m
	Possible Prizes		
N/A	$X = \{x, x', x'', \ldots\}$	$[\pi_X(x), \pi_X(x'), \ldots]$	N/A

*Answers to different questions are not generally independent. Typically, the joint probability measure $\pi \neq \pi_1 \cdots \pi_m \cdot \pi_X$.

Table 3: Representation of a cognitive state.

measure over $\Delta(\alpha)$ changes from π^0 to $\pi^{A_i} = \pi^0(\cdot|A_i)$.[9] We assume Bayesian updating here, which means that ex ante, before one knows what one will discover, an informational action determines a distribution over subjective judgments such that the expectation of this distribution equals the prior judgment. That is, by the law of total probability, $\sum_{A_i \in \mathcal{A}_i} \pi_i^0(A_i)\, \pi^{A_i} = \pi^0$. An informational action would decrease expected entropy because conditioning reduces entropy (see, e.g., Cover and Thomas, 1991, pg. 27). New information generates surprise (as formalized in the next section), which changes the attention weights too. Given the prior attention weight vector $\mathbf{w}^0$ based on salience and importance, we let $\mathbf{w}^{A_i}$ denote the new attention weight vector immediately after learning A_i, resulting from surprise at this discovery.

2.3 Preferences over (Distributions of) Cognitive States

The conventional theory of choice under risk assumes that a lottery over outcomes is evaluated according to its expected utility. Given that we may think of an informational action as creating a lottery over cognitive states, we make the natural assumptions leading to an expected utility representation in this new domain.

Independence Across Cognitive States

We assume that *there is a complete and transitive preference relation $\succeq$ on $\Delta(\mathcal{C})$ that is continuous (with respect to an appropriate topology)*[10] *and that satisfies independence*, so there exists a continuous expected utility representation u of $\succeq$ (von Neumann and Morgenstern, 1944).

The assumption here is that when information could put a person into one of many possible cognitive states, preference is consistent with valuing each possible cognitive state independently of any other cognitive states the person might have found herself in.

This might seem to imply that the utility of a state of uncertain knowledge is equal to the expected utility of each of the possible beliefs – e.g., that being uncertain of whether the object of my desire reciprocates my affections provides the same utility as the sum of probabilities times the utilities associated with the possible outcome belief states. It need not, because (as we discuss in detail below) obtaining the information, and indeed the specific information one obtains, is likely to affect one's attention weights. Such a change in attention can encourage or discourage a decision maker from resolving uncertainty, depending on whether the news that will be revealed is expected to be good or bad.

2.4 Choosing Between Sequences of Actions

The discovery of information following an initial action can change the availability or desirability of subsequent actions. For example, the information in a college professor's teaching ratings could help her decide whether to enroll in a teacher improvement class. A sequence of actions can be analyzed with the convention that an action operator passes through a distribution over cognitive states.[11] Thus, we represent a sequence of actions s acting on a cognitive state $(\pi, \mathbf{w})$ as $s\cdot(\pi, \mathbf{w}) \in \Delta(\mathcal{C})$.

Choice from among a set of sequences of actions $\mathcal{S}$, where early actions may reveal information that will inform later actions, is represented as utility maximization: a sequence $s^* \in \mathcal{S}$ may be chosen by a decision maker in the cognitive state $(\pi, \mathbf{w})$ if $s^* \in \arg\max_{s \in \mathcal{S}} u\,(s\cdot(\pi, \mathbf{w}))$. We find it useful to define a utility function over cognitive states, contingent on the set of sequences of actions that may subsequently be chosen:

$$U(\pi, \mathbf{w} \,|\, \mathcal{S}) = \max_{s \in \mathcal{S}} u\,(s\cdot(\pi, \mathbf{w}))\,. \tag{1}$$

In the example of the professor's teaching ratings, the set of available subsequent actions is to enroll in the teacher improvement class or not to enroll in the class. Looking at the ratings resolves a lottery over cognitive states, each of which having utility that is conditional on making the optimal choice of one of these subsequent actions.

We define the desirability of a sequence of actions s in cognitive state $(\pi, \mathbf{w})$ as $D(s \,|\, \pi, \mathbf{w}) = u\,(s\cdot(\pi, \mathbf{w})) - u\,(\pi, \mathbf{w})$.[12] Desirability is simply marginal utility relative to the trivial 'action' of doing nothing.

3. PSYCHOLOGICAL INSIGHTS

In this section we introduce a number of specific psychological insights that lead us to specify a utility function that generates a wide range of testable predictions concerning informational phenomena. These insights help us characterize the factors that influence the level of attention paid to a question as well as to identify distinctly the valence of beliefs and the desire for clarity.

[9] We thus denote a belief with complete certainty in $\mathbf{A}\times x$ as $\pi^{\mathbf{A}\times x}$.

[10] The induced topology on $\mathcal{C}$ (derived from the order topology on $\Delta(\mathcal{C})$) should be a refinement of the order topology on $\mathcal{C}$ (see Nielsen, 1984).

[11] Analogous to the standard assumption in decision under risk, the model assumes reduction of compound distributions over cognitive states. This does not imply the traditional reduction of compound lotteries.

[12] The degenerate distributions in $\Delta(\mathcal{C})$ correspond to individual states of knowledge. With the standard abuse of notation, we refer to the utility of the degenerate distribution on $(\pi, \mathbf{w}) \in \mathcal{C}$ as $u(\pi, \mathbf{w})$.

3.1 Attention

Neuroeconomic research indicates that attention shapes preference (Fehr and Rangel, 2011). Attention weights in our model specify how much a person is thinking about particular beliefs and, in turn, how much those beliefs directly impact utility. We may think of beliefs as having intrinsic value, which is then amplified by these attention weights. Our model (assuming monotonicity with respect to attention weights, as described in Section 4) provides a natural distinction between beliefs that have positive or negative intrinsic value: beliefs are positive specifically when more attention enhances utility and are negative in the opposite case. That is, a person likes thinking about (i.e., putting more attention weight on) *positive beliefs* and does not like thinking about *negative beliefs*.

Here we formalize the concepts of *importance*, *salience*, and *surprise*, all of which, we assume, contribute to attention weight. The importance γ_i of a question Q_i reflects the degree to which one's utility depends on the answer. Thus, for example, for an egocentric, but insecure, individual, the question, "Do other people like me?" is likely to be of great importance because the answer matters to the individual. Salience, distinctly, reflects the degree to which a particular context highlights the question. If, for example, an individual hears that another person was talking about her (with no further details), the question of whether the comments were favorable or not will become highly salient. We denote the salience of question Q_i as $\sigma_i \in \mathbb{R}_+$. Finally, surprise is a factor that reflects the dependence of attention on the dynamics of information revelation, and specifically on the degree to which receiving new information changes one's beliefs. If, having believed that she was generally well-liked, our individual were to discover that the comments about her were actually unfavorable, the discovery, necessitating a radical change in her belief, would be quite surprising (and, as we presently assume, would increase her attention to the question). We denote the surprise associated with a revised belief about question Q_i as δ_i. We assume that *the attention w_i on an activated question Q_i is a strictly increasing function of this question's importance γ_i, its salience σ_i, and the surprise δ_i associated with it.*

Importance

The importance of a question depends on the spread of the utilities associated with the different answers to that question. The degree to which an individual's utility varies with the answers to a question depends both on the magnitude of the utility function and on the perceived likelihood of different answers. Continuing with the example of the question of how well-liked an individual is, one could distinguish two relevant traits: egocentrism – the degree to which the individual *cares* about being well-liked; and insecurity – the dispersion of the individual's subjective probability distribution across possible answers. By our definition of the concept, importance should be positively related to both traits.

Given a particular prior subjective probability measure π^0 and a set $\mathcal{S}$ of sequences of actions available to the decision maker, the importance γ_i of question Q_i is a function (only) of the likelihood of possible answers and the utilities associated with these answers, captured as

$$\gamma_i = \phi\left(\left\langle \pi_i^0(A_i),\, U\big(\pi^{A_i}, \mathbf{w}^{A_i} \,|\, \mathcal{S}\big)\right\rangle_{A_i \in \operatorname{supp}\left(\pi_i^0\right)}\right)$$

where U is the utility function defined in Equation (1). Without specifying the precise form of this function ϕ, we assume only that *it (i.e., importance) increases with mean-preserving spreads of the (subjective) distribution of utilities that would result from different answers to the question*, and that *it is invariant with respect to constant shifts of utility.* Thus, a question is important to the extent that one's utility depends on the answer. Raising the stakes increases importance. On the other hand, if an answer is known with certainty, then by this definition nothing is at stake, so the underlying question is no longer important. While acquiring information will affect the importance of the questions being addressed, it takes time to adapt to news, so there should be some delay. We assume that *the importance of a question is updated only when the new information is incorporated into a new default subjective probability measure.*

Our definition of importance is, admittedly, circular. Importance depends on utility, which in turn depends on the attention weight, but importance also contributes to attention weight. There is, likely, some psychological realism to this circularity which captures the dynamic processes giving rise to obsession: attention to a question raises its importance, and the elevated importance gives rise to intensified attention. If we assume that these processes unfold instantaneously, then importance (and, in turn, attention weight and utility) will be a fixed point of this composition of functions. We can make simple comparisons of importance without going to the trouble of specifying precise values.

Salience

The salience of a question depends on a variety of exogenous contextual factors. For example, a question could be salient if it has recently come up in conversation (i.e., it has been primed) or if other aspects of the environment remind an individual about it. Alternatively, a question could be more salient to an individual if the answer is, in principle, knowable, and even more so if other people around her know the answer but she does not.

Often a question may be salient despite being unimportant. Continuing the prior example, even if an individual deems others' perceptions of her as unimportant, the question of her popularity might nonetheless be highly salient if the individual was asked, "Do you know what x thinks of you?" Conversely, there are myriad questions that are important by the definition just provided, but which lack salience. There might be numerous people whose opinion of us we would care about and be unsure of, but unless something raises the issue in our mind, we are unlikely to focus on it. It seems natural to think that some degree of salience is a necessary, and sufficient, condition for attention (while some degree of importance is not). Thus, we assume that a question Q_i is activated (i.e., has strictly positive attention weight $w_i > 0$) if and only if it has positive salience $\sigma_i > 0$. Further, we assume that *attention weight w_i has strictly increasing differences (i.e., a positive cross-partial derivative, if we assume differentiability) in* (γ_i, σ_i). That is, an increase in importance produces a greater increase in attention for a more salient question.

Surprise

The third factor that we posit influences attention is the surprise one experiences upon acquiring new information. Surprise reflects the degree to which new information changes existing beliefs. A natural measure of surprise was proposed in a theoretical paper by Baldi (2002) and, in an empirical follow-up investigation (Itti and Baldi, 2009), shown to predict the level of attention paid to information. Incorporating the insights from this line of research, we assume that *when the answer to a particular question Q_j is learned, thereby contributing information about the answers to associated questions and causing their subjective probabilities to be updated, the degree of surprise associated with a new belief about question Q_i can be defined as the Kullback-Leibler divergence of $\pi_i^{A_j}$ against the prior π_i^0:*

$$\delta_i(\pi_i^{A_j}||\pi_i^0) = \sum_{A_i \in \mathcal{A}_i} \pi_i^{A_j}(A_i) \log \frac{\pi_i^{A_j}(A_i)}{\pi_i^0(A_i)}.$$

Surprise is positive with any new information, and is greatest when one learns the most unexpected answer with certainty. However, the feeling of surprise is not permanent. We assume that *when the decision maker adapts and gets used to this new knowledge (formally, when the default subjective probability measure is reset), it is no longer surprising.*

The Belief Resolution Effect

The impact of new information on attention is greatest when uncertainty about a question is resolved completely. Surprise immediately spikes, but in the long run fades, and the underlying question becomes unimportant because, with the answer known, there is no longer a range of possible answers. Taken together, these factors create a pattern of change in attention weight following the discovery of a definitive answer, what we call the *belief resolution effect* – when an answer is learned with certainty, there is an immediate boost in attention weight on it, but over time this attention weight falls to a lower level. Specifically, when the decision maker adapts and the certain belief is incorporated into the default subjective probability measure, the question then receives less attention. It is as if the brain recognizes that because a question has been answered, it can move on to other questions that have yet to be addressed. Janis (1958) recognized the belief resolution effect when he observed that surgical patients getting information about their upcoming procedures initially worry more about the surgery but subsequently experience less anxiety.

3.2 Valence and Clarity

It is useful to distinguish two sources of a belief's intrinsic value: *valence* and *clarity*. Valence refers to the value attached to answers to questions. To illustrate the concept of valence, we return to the example of a professor's belief about her teaching ability. Being a good (or bad) teacher carries intrinsically positive (or, respectively, negative) valence. Clarity refers to preferences between degrees of certainty, independent of the answers one is certain of. We assume that, *ceteris paribus, people prefer to have greater clarity (i.e., less uncertainty or more definitive subjective beliefs).* The aversion that people feel towards uncertainty is reflected in neural responses in the anterior cingulate cortex, the insula and the amygdala (Hirsh and Inzlicht, 2008; Sarinopoulos et al., 2010). It manifests in physiological responses as well. Subjects who know to expect an electric shock, but who are uncertain whether it will be mild or intense, show more fear – they sweat more profusely, and their hearts beat faster – than subjects who know for sure that an intense shock awaits (Arntz et al., 1992).

When valence and clarity pull in opposite directions, it may be the case that people prefer a certain answer to a subjective belief that dominates it on valence or that people prefer uncertainty when it leaves space for better answers. While the preference for clarity violates Savage's (1954) sure-thing principle, we do assume a weaker version of it:

One-Sided Sure-Thing Principle

For any $\pi \in \Delta(\alpha)$, let $\mathrm{supp}(\pi) \subseteq \alpha$ denote the support of π. If for all $\mathbf{A}\times x \in \mathrm{supp}(\pi)$ we have $u(\pi', \mathbf{w}) \geq u(\pi^{\mathbf{A}\times x}, \mathbf{w})$, then $u(\pi', \mathbf{w}) \geq u(\pi, \mathbf{w})$, with the latter inequality strict whenever there exist $\mathbf{A}' \times x'$ and $\mathbf{A}'' \times x'' \in \mathrm{supp}(\pi)$ such that $\mathbf{A}' \neq \mathbf{A}''$.

The one-sided sure-thing principle asserts that people always prefer a certain answer to uncertainty amongst answers that all have valences no better than the certain answer (holding attention weight constant).

A Measure of Uncertainty

The assumption of a preference for clarity means that there is a preference for less uncertain subjective beliefs. A useful measure of the uncertainty about a particular question is the entropy of the subjective probability distribution over answers (Shannon, 1948). The entropy of a subjective (marginal) probability π_i is $H(\pi_i) = -\sum_{A_i \in \mathcal{A}_i} \pi_i(A_i) \log \pi_i(A_i)$ (with the convention that $0 \log 0 = 0$).[13] At one extreme, entropy is high when there are many equally likely possible answers; at the other extreme, there is minimal entropy of 0 when a single answer is known for sure.

3.3 A Specific Utility Function

To make precise predictions about preferences for (or to avoid) information, we consider a specific utility function incorporating the preference for clarity and the role of attention weights:

$$u(\pi, \mathbf{w}) = \sum_{x \in X} \pi_X(x) v_X(x) + \sum_{i=1}^{m} w_i \left(\sum_{A_i \in \mathcal{A}_i} \pi_i(A_i) v_i(A_i) \ - H(\pi_i) \right). \quad (2)$$

We represent the value of prize x as $v_X(x)$ and the valence of answer A_i as $v_i(A_i)$. We now describe properties (some quite strong and almost certainly not always satisfied) that characterize (and necessarily imply) this utility function (see Theorem 4.2 below).

4. CHARACTERIZATION OF THE UTILITY FUNCTION

4.1 Properties

The utility function in Equation (2) satisfies the following seven properties.

Independence Across Prizes

In Section 2 we assumed independence across cognitive states. Independence might extend, as in traditional models, to material outcomes, holding beliefs constant.

[13]The base of the logarithm in the entropy formula is arbitrary and amounts to a normalization parameter.

P1. Holding the rest of the cognitive state constant, the preference relation satisfies independence across prizes if $u(\pi^{\mathbf{A}}, \mathbf{w}) = \sum_{x \in X} \pi_X^{\mathbf{A}}(x)\, u(\pi^{\mathbf{A} \times x}, \mathbf{w})$.

Property (P1) implies belief-dependent expected utility over lotteries that are independent of beliefs about the world. If we also were to assume belief-independent utility for prizes, then we would gain the ability to reduce compound lotteries consisting of horse races as well as roulette lotteries (Anscombe and Aumann, 1963) to single-stage lotteries. However, we believe it is often the case that utility is belief-dependent. We might say that a decision maker often has a horse in the race.

Separability Between Questions

Additive separability of utility between questions means that a person can place a value on a belief about a given question without needing to consider beliefs about other questions.

P2. A utility function satisfies additive separability between questions if $u(\pi, \mathbf{w}) = u_X(\pi_X) + \sum_{i=1}^{m} u_i(\pi_i, w_i)$.[14]

Property (P2) may seem quite strong because we can imagine representations of sensible preferences that are not additively separable. For example, the value of a belief about whether a car on sale has a warranty intuitively could depend on the cost of the car in the first place (not to mention one's desire for a new car, one's estimation of the costs of car repairs, etc.). However, we may be able to represent these preferences as separable after all. We might suppose that these beliefs do have separable values but that they correlate with some other highly valued belief, perhaps about how good a deal one can get on the car. That is, while intuition tells us that the value of beliefs about different questions (e.g., "does she like me?" and "does she have a boyfriend?") is often interdependent, this dependence may be mediated by the existence of additional questions (e.g., "will she go out with me?"), beliefs about which may be mutually dependent, but independently valued.

Monotonicity with respect to Attention Weights

Preferences satisfy the property of monotonicity with respect to attention weights if whenever increasing attention on a given belief enhances (or diminishes) utility, it will do so regardless of the absolute level of attention weight. At a psychological level, the interpretation of this monotonicity property is that when a belief is positive, more attention to it is always better, and when a belief is negative, more attention is always worse. In fact, the property provides a natural *definition* of whether a belief is positive or negative.

P3. Preferences satisfy monotonicty with respect to attention weights if for any $\mathbf{w}$, $\hat{\mathbf{w}}$, and $\mathbf{w} \in \mathbb{R}_+^m$ such that $w_i = \hat{w}_i = w_i$ for all $i \neq j$ and $w_j > \hat{w}_j > w_j$, we have $u(\pi, \hat{\mathbf{w}}) \geq u(\pi, \mathbf{w})$ if and only if $u(\pi, \mathbf{w}) \geq u(\pi, \hat{\mathbf{w}})$, with equality on one side implying equality on the other, for all $\pi \in \Delta(\alpha)$.

In the case that these inequalities hold strictly, we say that π_j, the belief about question Q_j, is a *positive belief.* If they hold as equalities, we say that π_j is a *neutral belief.* And, in the case that the inequalities hold in the reverse direction, then π_j is a *negative belief.*

Linearity with respect to Attention Weights

The next property describes how changing the attention on a belief impacts utility. For any given attention weight, the marginal utility of a change in belief depends on what those beliefs are and how much the individual values them. The property of linearity with respect to attention weights means that, in general, the marginal utility associated with such a change in belief (assuming the utility of this belief is separable) is proportional to the attention on that belief.

P4. When the utility of question Q_i is separable, linearity with respect to attention weights is satisfied if for any w_i and $\hat{w}_i \in \mathbb{R}_+$ and π_i' and $\pi_i'' \in \Delta(\mathcal{A}_i)$, we have

$$u_i(\pi_i', \hat{w}_i) - u_i(\pi_i'', \hat{w}_i) = \frac{\hat{w}_i}{w_i} \left(u_i(\pi_i', w_i) - u_i(\pi_i'', w_i) \right).$$

Property (P4) allows us, in the case of separable utility, to assign an intrinsic value v to beliefs such that $u_i(\pi_i', w_i) - u_i(\pi_i'', w_i) = w_i \left(v_i(\pi_i') - v_i(\pi_i'') \right)$. We abuse notation by referring to the valence of answer A_i as $v_i(A_i)$, with it being defined here as the intrinsic value v_i of belief with certainty in A_i. We have taken the liberty of specifying a precise relationship between attention weights and utility as a convenient simplification; it should be noncontroversial because we do not claim to have a cardinal measure of attention weight.

Label Independence

Intuitively, the value of a belief should depend on how an individual values the possible answers and on how probable each of these answers is, and these factors (controlling for attention weight of course) should be sufficient to determine the utility of any (uncertain) belief. In particular, the value of a belief should not depend on how the question or the answers are labeled.

P5. Label independence is satisfied if, when the utility of questions Q_i and Q_j are separable, a bijection $\tau : \mathcal{A}_i \to \mathcal{A}_j$, such that $v_i(A_i) = v_j(\tau(A_i))$ and $\pi_i(A_i) = \pi_j(\tau(A_i))$, implies that $v_i(\pi_i) = v_j(\pi_j)$.

Reduction of Compound Questions

The intuition behind the assumption of label independence also seems to suggest that the utility of a belief perhaps should not depend on the way the question giving rise to the belief is asked, i.e., on whether a complicated question is broken up into pieces. We should recall, however, that the activation of a particular question directs attention to the belief about this question. Thus, in general, the utility of a belief will not be invariant to the question being asked. Still, it may be the case that utility remains invariant when a compound question is broken into parts as long as the attention on each part is weighted properly. If utility remains invariant upon setting attention weights on conditional questions to be proportional to the subjective probabilities of the hypothetical conditions, then we say that the utility function satisfies the reduction of compound questions property. Figure 1 demonstrates the reduction of a compound question with appropriate attention weights on each subquestion.

[14] A subset of questions $\tilde{\mathcal{Q}} \subset \mathcal{Q}$ can also be separable, in which case $u(\pi, \mathbf{w}) = \sum_{i: Q_i \in \tilde{\mathcal{Q}}} u_i(\pi_i, w_i) + u_{-\tilde{\mathcal{Q}}}(\pi_{-\tilde{\mathcal{Q}}}, \mathbf{w}_{-\tilde{\mathcal{Q}}})$ where $\pi_{-\tilde{\mathcal{Q}}}$ is the marginal distribution over answers to the remaining questions and prizes and the vector $\mathbf{w}_{-\tilde{\mathcal{Q}}}$ contains the remaining components of $\mathbf{w}$.

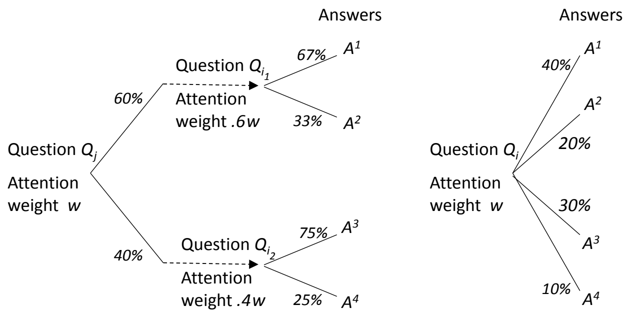

Figure 1: Decomposition of a compound question.

P6. A separable utility function satisfies the reduction of compound questions property if whenever there is a partition ζ of the answers $\mathcal{A}_i$ (to question Q_i) into $\zeta = \{\mathcal{A}_{i_1}, \ldots, \mathcal{A}_{i_n}\}$ and a bijection $\tau : \zeta \to \mathcal{A}_j$ into the answers to some question Q_j such that for any $h \in [1, n]$ and any $A_i \in \mathcal{A}_{i_h}$,

$$v_i(A_i) = v_j(\tau(\mathcal{A}_{i_h})) + v_{i_h}(A_i)$$

and

$$\pi_i(A_i) = \pi_j(\tau(\mathcal{A}_{i_h})) \cdot \pi_{i_h}(A_i),$$

it follows that

$$u_i(\pi_i, \omega) \;=\; u_j(\pi_j, \omega) \;+ \sum_{h=1}^{n} u_{i_h}(\pi_{i_h}, \, \pi_j(\tau(\mathcal{A}_{i_h})) \cdot \omega).$$

Ruling Out Unlikely Answers Increases Clarity

A final property operationalizes the preference for clarity. Controlling for the valence of one's beliefs, by considering situations in which one is indifferent between different possible answers to a question, there should be a universal aversion to being uncertain about the answer to an activated question. As a building block toward quantifying the uncertainty in a subjective belief, we assert here that when an unlikely (and equally attractive) answer is ruled out, uncertainty decreases (and thus the utility of that uncertain belief increases).

P7. Ruling out unlikely answers increases clarity if, when the utility of question Q_i is separable and all answers to this question have the same valence, i.e. $v_i(A_i) = v_i(A_i')$ for all A_i and $A_i' \in \mathcal{A}_i$, then for any π where without loss of generality $\pi_i(A_i^h)$ is weakly decreasing in h and for any π' such that $\pi_i'(A_i^h) \geq \pi_i(A_i^h)$ for all $h \in [1, \bar{h}]$ (with at least one inequality strict) and $\pi_i'(A_i^h) = 0$ for all $h > \bar{h}$, for some $\bar{h}$, we consequently have $v_i(\pi_i') > v_i(\pi_i)$.

4.2 Utility Representation Theorem

If the properties P1-P7 are satisfied, then

$$u(\pi, \mathbf{w}) = \sum_{x \in X} \pi_X(x) v_X(x) + \sum_{i=1}^{m} w_i \left(\sum_{A_i \in \mathcal{A}_i} \pi_i(A_i) v_i(A_i) \; - H(\pi_i) \right).$$

PROOF. Linearity with respect to attention weights allows us to pull an attention weight on question Q_i outside of the utility $u_i(\pi_i, w_i) = w_i v_i(\pi_i)$ (using a neutral belief to calibrate v_i). A partition of $\mathcal{A}_i$ into singletons $\mathcal{A}_{i_h}$ such that $v_i(A_i) = v_{i_h}(A_i)$ allows us, by reduction of the compound question, to determine that the function $F(\pi_i) = v_i(\pi_i) - \sum_{A_i \in \mathcal{A}_i} \pi_i(A_i) v_i(A_i)$ does not depend on $v_i(A_i)$ for any $A_i \in \mathcal{A}_i$. Moreover, $-F(\cdot)$ satisfies Shannon's (1948) axioms (continuity, increasing in the number of equiprobable answers, and reduction of compound questions) characterizing the entropy function $H(\pi_i) = -\sum_{A_i \in \mathcal{A}_i} \pi_i(A_i) \, \log \pi_i(A_i)$. $\square$

5. INFORMATION ACQUISITION AND AVOIDANCE

We can apply our utility function to decisions about information acquisition or avoidance. We develop our analysis in a companion paper (Golman and Loewenstein, 2015a), and we provide a broad outline here of its implications. The desire for information, in our model, can be decomposed into three distinct motives: recognition of the instrumental value of the information; curiosity to fill the information gap(s); and motivated attention to think more or less about what could be discovered. The instrumental value of information arises from its impact on subsequent actions. As in the standard account of informational preferences, it is defined as the difference between the expected utility of subsequent actions conditional on having the information and the utility expected in the absence of the information. Curiosity arises from the expected reduction in uncertainty upon acquiring information. It is defined as the expected utility of revised beliefs, given prior levels of attention. The magnitude of curiosity depends on the attention devoted to each information gap that stands to be addressed. Motivated attention arises from the surprise upon acquiring information. It is defined as the expected utility from increased attention on whatever happens to be discovered, conditioning on all possible outcomes. Motivated attention is a motive to acquire information that's expected to be good and to avoid information that's expected to be bad.

Putting the three motives together, our model makes many predictions about when, and the degree to which, information will be sought or avoided. When anticipated answers are neutral or even potentially positive, information should be sought. The strength of the desire for this information should increase with the number of attention gaps that can be addressed, the attention paid to them, and the valence of the possible outcomes. However, when anticipated outcomes are sufficiently negative, information would be avoided. This "ostrich effect" when anticipating bad outcomes is consistent with a growing body of empirical evidence (see, e.g., Karlsson et al., 2009; Eil and Rao, 2011). In addition, the belief-resolution effect in our model leads to a novel prediction: individuals who discount the future less should be less likely to exhibit the ostrich effect and more likely to acquire information despite anticipated bad news.

6. RISK AND AMBIGUITY PREFERENCE

Section 5 outlines how the model we have developed allows us to describe a desire to acquire or to avoid information. We can apply this same model to an entirely new domain: preferences about wagers that depend on missing information. Risk and ambiguity aversion are complex topics, and we develop these applications in depth in a companion paper (Golman and Loewenstein, 2015b). Here, we provide a broad outline of the model's implications in this domain.

Decision making under risk and under ambiguity both expose decision makers to information gaps. Imagine a choice between a gamble and a sure thing. Deciding to play the gamble naturally focuses attention on the question: what will be the outcome of the gamble? Of course, deciding to not play the gamble does not stop an individual from paying some attention to the same question (or, if not choosing the gamble means it will not be played out, the related question: what *would have been* the outcome of the gamble?) but playing the gamble makes the question more important, and that brings about an increase in the attention weight on the question. If the individual is aware of this effect, which is natural to assume, then whether it encourages risk taking or risk aversion will depend on a second factor: whether thinking about the information gap is pleasurable or aversive. When thinking about the missing information is pleasurable, then the individual will be motivated to increase attention on the question, which entails betting on it. Conversely, when thinking about the missing information is aversive, the individual will prefer to not bet on it. This may help to explain why, for example, people generally prefer to bet on their home teams rather than on other teams, especially in comparison to the home team's opponent.

Decision making involving uncertainties that are ambiguous is similar to the case with known risks, but with an additional wrinkle: with ambiguity, there are additional information gaps. In a choice between a sure thing and an ambiguous gamble, for example, a second relevant question (in addition to the one above about the outcome of the gamble) is: what is the probability of winning with the ambiguous gamble? (And there may be additional relevant questions that could inform someone about this probability, so even a Bayesian capable of making subjective probability judgments would be exposed to these information gaps.) Again, betting on the ambiguous gamble makes these questions more important and thus will increase the attention weight on them. So, desire to play the gamble will be increasing with the degree to which thinking about the gamble is pleasurable. To the extent that abstract uncertainties are not pleasurable to think about, this model provides a novel account of standard demonstrations of ambiguity aversion, including those first generated by Ellsberg (1961) in his seminal paper on the topic.

7. DESIRE FOR WISDOM

Our utility model can be used to describe preferences between knowing and not knowing. But another comparison is also of interest, albeit harder to investigate empirically – the difference between awareness and unawareness. While we cannot easily give a person the choice whether or not to become aware of a question, we can at least introspect. We might posit that awareness of meaningful questions is a source of utility. Equation (2), the utility function which represents preferences between cognitive states given a fixed set of activated questions $\mathcal{Q}$, might be augmented with a term $v_{\mathcal{Q}}(\mathcal{Q})$ capturing the intrinsic value of awareness of particular issues.

Wisdom, the combination of awareness and clarity,[15] is, or at least tends to be, preferable to ignorance. We of course

Question	Answer	Belief		
Latent	–	Unawareness	Awareness	Wisdom
Activated	Unknown	Uncertainty	Clarity	
	Known	Certainty		

Table 4: Wisdom, the combination of awareness and clarity.

must allow exceptions if we are serious that beliefs have valence that may be negative. The popular adage that "ignorance is bliss" expresses concern for the negative beliefs that awareness may entail. However, in many natural situations, a person may reasonably anticipate that newfound awareness will bring about neutral or even positive beliefs. In such contexts, information and awareness may be simultaneously acquired. For example, a bird-watcher typically would strictly prefer to learn the name of a previously unnoticed songbird rather than to remain unaware of its existence. Curiosity is behind the desire to catch the name upon becoming aware of the bird's existence, even though the particular name does not really matter, but utility from awareness implies that opening, and then immediately closing, an aversive information gap need not be zero sum. Rather, discovering the new bird's name, acquiring both the question and the definitive answer, produces a net positive utility gain, which is what we designate, in the context of our model, the *utility of wisdom.* We find the desire for wisdom in individuals' varied pursuits of insight and expertise, from a naturalist's passion for identifying flora and fauna to a fan's thirst for new baseball statistics or a connoisseur's discriminating taste for wine.[16]

Aristotle in 350 B.C. asserted, "All men by nature desire to know." John Stuart Mill agreed, in his classic *Utilitarianism*, arguing that, "It is better to be a human being dissatisfied than a pig satisfied; better to be Socrates dissatisfied than a fool satisfied." We too assert that knowledge can be a very real source of utility. A perspective that information derives value solely from its ability to yield material consumption fails to appreciate the most profound benefits provided by information, the knowledge and wisdom it confers.

8. REFERENCES

Abelson, R. (1986). Beliefs are Like Possessions. *Journal for the Theory of Social Behavior* 16 (3), 223-250.

Anscombe, F., Aumann, R. (1963). A Definition of Subjective Probability, *Annals of Mathematical Statistics* 34, 199-205.

Arntz, A., Van Eck, M., de Jong, P. (1992). Unpredictable Sudden Increases in Intensity of Pain and Acquired Fear. *Journal of Psychophysiology* 6 (1), 54-64.

Asch, D., Patton, J., Hershey, J. (1990). Knowing for the Sake of Knowing: The Value of Prognostic Information. *Medical Decision Making* 10 (1), 47-57.

Aumann, R. (1976). Agreeing to Disagree. *Annals of Statistics* 4, 1236-1239.

Baldi, P. (2002). A Computational Theory of Surprise. In *Information, Coding, and Mathematics*, M. Blaum, P. Farrell, H. van Tilborg (Eds.) Norwell, MA: Kluwer.

[15] We are aware that this may not be the most common usage of the word *wisdom*, but the distinction between knowledge acquired from a state of uncertainty and knowledge acquired from a state of unawareness is rarely made explicit. The term, "wisdom" seems to adequately capture this distinction if we think of a wise man or woman as not only having the right answers, but also asking the right questions.

[16] Lab studies also find that people prefer environments which seem to stimulate new questions and promise to provide relevant information (Kaplan, 1992).

Caplin, A., Leahy, J. (2001). Psychological Expected Utility Theory And Anticipatory Feelings. *Quarterly Journal of Economics* 116 (1) 55-79. view of Economic Studies 67, 529- 544.

Cover, T., Thomas, J. (1991). *Elements of Information Theory*. New York: Wiley.

Dekel, E., Lipman, B., Rustichini, A. (1998). Standard State-Space Models Preclude Unawareness. *Econometrica* 66 (1), 159-173.

Dillenberger, D. (2010). Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior. *Econometrica* 78 (6), 1973-2004.

Eil, D., Rao, J. (2011). The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself. *American Economic Journal: Microeconomics* 3, 114-138.

Eliaz, K., Spiegler, R. (2006). Can Anticipatory Feelings Explain Anomalous Choices of Information Sources? *Games and Economic Behavior* 56 (1), 87-104.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75, 643-699.

Fehr, E., Rangel, A. (2011). Neuroeconomic Foundations of Economic Choice – Recent Advances. *Journal of Economic Perspectives* 25 (4), 3-30.

Gabaix, X., Laibson, D., Moloche, G., Weinberg, S. (2006). Costly Information Acquisition: Experimental Analysis of a Boundedly Rational Model. *American Economic Review* 96 (4), 1043-1068.

Geanakoplos, J., Pearce, D., Stacchetti, E. (1989). Psychological games and sequential rationality. *Games and Economic Behavior* 1, 60-79.

Golman, R., Loewenstein, G. (2015a). Curiosity, Information Gaps, and the Utility of Knowledge. Working Paper.

Golman, R., Loewenstein, G. (2015b). Information Gaps for Risk and Ambiguity. Working Paper.

Grant, S., Kajii, A., Polak, B. (1998). Intrinsic Preference for Information. *Journal of Economic Theory* 83, 233-259.

Heifetz, A., Meier, M., Schipper, B. (2006). Interactive Unawareness. *Journal of Economic Theory* 130 (1), 78-94.

Hirsh, J., Inzlicht, M. (2008). The Devil You Know: Neuroticism Predicts Neural Response to Uncertainty. *Psychological Science* 19 (10), 962-967.

Hirshleifer, J., Riley, J. (1979). The Analytics of Uncertainty and Information – An Expository Survey. *Journal of Economic Literature* 17 (4), 1375-1421.

Itti, L., Baldi, P. (2009). Bayesian Surprise Attracts Human Attention. *Vision Research* 49 (10), 1295-1306.

Janis, I. (1958). *Psychological Stress: Psychoanalytic and Behavioral Studies of Surgical Patients*. New York: Wiley.

Kadane, J., Schervish, M., Seidenfeld, T. (2008). Is Ignorance Bliss? *Journal of Philosophy* 105 (1), 5-36.

Kang, M.J., Hsu, M., Krajbich, I., Loewenstein, G., McClure, S., Wang, J., Camerer, C. (2009). The Wick in the Candle of Learning: Epistemic Curiosity Activates Reward Circuitry and Enhances Memory. *Psychological Science* 20 (8), 963-973.

Kaplan, S. (1992). Environmental preference in a knowledge-seeking, knowledge-using organism. In *The Adapted Mind*, J. H. Barkow, L. Cosmides, and J. Tooby (Eds.), New York: Oxford UP.

Karlsson, N., Loewenstein, G., McCafferty, J. (2004). The Economics of Meaning. *Nordic Journal of Political Economy* 30 (1), 61-75.

Karlsson, N., Loewenstein, G., Seppi, D. (2009). The Ostrich Effect: Selective Attention to Information. *Journal of Risk and Uncertainty* 38 (2), 95-115.

Karni, Edi, and Marie-Louise Vierø. (2013). “Reverse Bayesianism”: A Choice-Based Theory of Growing Awareness. *American Economic Review* 103 (7), 2790-2810.

Klibanoff, P., Marinacci, M., Mukerji, S. (2005). A Smooth Model of Decision Making under Ambiguity. *Econometrica* 73, 1849-1892.

Köszegi, B. (2010). Utility from Anticipation and Personal Equilibrium. *Economic Theory* 44 (3), 415-444.

Kreps, D., Porteus, E. (1978). Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica* 46 (1), 185-200.

Li, J. (2008). A Note on Unawareness and Zero Probability. PIER Working Paper No. 08-022.

Locke, E., Latham, G. (1990). *A Theory of Goal Setting and Task Performance*. Englewood Cliffs, NJ: Prentice Hall.

Loewenstein, G. (1994). The Psychology of curiosity: A review and reinterpretation. *Psychological Bulletin* 116 (1), 75-98.

Loewenstein, G. (1999). Because it is there: The challenge of mountaineering... for utility theory. *Kyklos* 52, 315-344.

Modica, S., Rustichini, A. (1994). Awareness and Partitional Information Structures. *Theory and Decision* 37 (1), 107-124.

Modica, S., Rustichini, A. (1999). Unawareness and Partitional Information Structures. *Games and Economic Behavior* 27 (2), 265-298.

Miller, G., Galanter, E., Pribram, K.H. (1960). *Plans and the Structure of Behavior*. New York: Holt.

Nielsen, L. (1984). Unbounded Expected Utility and Continuity. *Mathematical Social Sciences* 8, 201-216.

Rabin, M. (2000). Risk Aversion and Expected-utility Theory: A Calibration Theorem. *Econometrica* 68, 1281-92.

Ritov, I., Baron, J. (1990). Reluctance to vaccinate: Omission bias and ambiguity. *Journal of Behavioral Decision Making* 3 (4), 263-277.

Sarinopoulos, I., Grupe, D., Mackiewicz, K., Herrington, J., Lor, M., Steege, E., Nitschke, J. (2010). Uncertainty During Anticipation Modulates Neural Responses to Aversion in Human Insula and Amygdala. *Cerebral Cortex* 20 (4), 929-940.

Savage, L. (1954). *The Foundations of Statistics*. New York: Wiley.

Schelling, T. (1987). The Mind as a Consuming Organ. In *The Multiple Self*, J. Elster (Ed.), Cambridge: Cambridge UP, 177-196.

Shannon, C. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal* 27, 379-423.

Tasoff, J., Madarász, K. (2009). A Model of Attention and Anticipation. Working Paper.

von Neumann, J., Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton: Princeton UP.

Wakker, P. (1988). Nonexpected Utility as Aversion of Information. *Journal of Behavioral Decision Making* 1 (3), 169-175.

Yariv, L. (2001). Believe and Let Believe: Axiomatic Foundations for Belief Dependent Utility Functionals. Cowles Foundation Discussion Paper No. 1344.