A Psychologically Plausible Goal-Based Utility Function

T. Grandon Gill
University of South Florida, Tampa, Florida, USA

ggill@coba.usf.edu

Abstract

Utility is a concept that is central to economics, finance, and decision theory. In the informing sciences, it also plays a central role: that of explaining the client's motivation to be informed. Unfortunately, existing economic models of utility are plagued by many anomalies when applied to individual clients and also fail to incorporate many important features that are relevant to informing contexts, such as the desire for mastery and learning. The present paper draws upon goal setting theory and cognitive psychology to propose an alternative view of utility that is better adapted to informing situations. It further proposes that in a given task setting, an individual's utility function will evolve with experience, from satisfying loose generic goals into a highly compiled task progress monitoring function. During the course of this evolution, the client will pass through a phase where utility is heavily determined by progress towards specific task goals. The implications of this more psychologically plausible model of utility for informing systems are discussed.

Keywords: Utility, goal setting, motivation, learning, informing systems, preferences, altruism, mastery, attention, framing, anchoring.

Introduction

The notion of a utility function, that is to say a function that captures human preferences and thereby explains human choices, is widely used in many business-related fields. In economics, it plays a foundational role in the mathematical derivation of general equilibrium theory and often plays a role in macroeconomic policy models. In the decision sciences, it is central to rational decision models involving risk and time. In management and finance, it serves as an important underpinning to a wide range of models, such as those employed in agency theory and prospect theory.

Within the informing sciences, the concept of utility is no less important. In Cohen's (1999) sender-channel-client model, the concept of utility is central to understanding why a sender chooses to inform and why a client chooses to be informed. In this context, utility can manifest itself in a number of ways. From an economic perspective, the value of being informed can be described as the marginal value of information for the task associated with the informing system. The greater the value of the information, the greater the contribution informing makes to the client's utility. Informing relationships may also exist, however, in contexts where marginal values cannot be calculated, or where

they are irrelevant. Learning is often intrinsically motivated; we experience utility gains through the process of being informed, yet there is not necessarily an economic source that we can pinpoint to explain that utility. Neoclassical economic theory allows utility to include less tangible elements, such as personal sense of mastery, power, or the satisfaction of others (Simon, 1992). Incorporating such elements, however, makes the utility construct substantially less tractable for the purposes of generating formal economic theory.

Neoclassical economic utility has two largely unresolved aspects that make its use particularly problematic in the informing context. Even in its simplest form, a single variable and one time period, numerous departures from apparently rational behavior as predicted by neoclassical utility functions have been observed in experimental settings. Henceforth, we refer to these as anomalies. Many of these anomalies (e.g., framing, priming) represent behaviors where subtle aspects of how a decision task is presented—not affecting the underlying decision itself—result in changes to observed client preferences. Obviously, better understanding how such anomalies arise is critical if effective informing is to take place, since the presentation of task information is an integral part of most informing systems.

Even more troubling, from an informing perspective, is the degree to which learning is ignored in economic models of utility. Utility is presumed to guide our choices, yet we also know that our preferences may evolve as we gain expertise in a particular activity. Economists do posit a roll for changing tastes—although some (e.g., Becker, 1976) do their best to propose functions constructed so as to minimize the impact of such changes—but when our utility preferences change (e.g., as we gain expertise in performing a particular type of financial task), it is not convincing to assert that this change is solely a consequence of our changing taste for money. Even more to the point, how plausible is it to assume that preferences do not change as a consequence of the learning that invariably takes place over the course of an informing relationship?

The purpose of the present paper is to synthesize existing research into an alternative formulation of the utility function that may serve to address some of the anomalies and inconsistencies common to many existing utility functions, particularly in the informing context. The paper begins by summarizing existing models of utility and then reviews key findings of the goal-setting literature. It then formulates a utility model driven by individual goals, which is then expanded to incorporate learning. Finally, the paper considers how predictions of the goal-based model might differ from existing models and the implications of these differences for the informing sciences.

Utility Theory

Utility generally refers to an individual state that combines notions of satisfaction, usefulness, and rationality. In economics, utility is often presented as a function of various goods and services $(x_1, x_2, \ldots, x_n)$ consumed in a period along with resources saved $(s_1, s_2, \ldots, s_n)$, e.g.,

$$U(x_1, x_2, \ldots, x_n, s_1, s_2, \ldots, s_n)$$

Neoclassical economic theory is built upon the assumption that any rational economic agent can be modeled as a utility maximizer. Conceptually, we may view the utility maximization process as a choice process wherein a decision-maker starts with a set of saved resources supplemented by earnings, considers how those resources may change as a consequence of possible actions available to him or her, and then selects that action leading to the highest expected utility outcome (where the outcome of an action may be a set of possible outcomes where uncertainty is present). We represent this model in Figure 1.
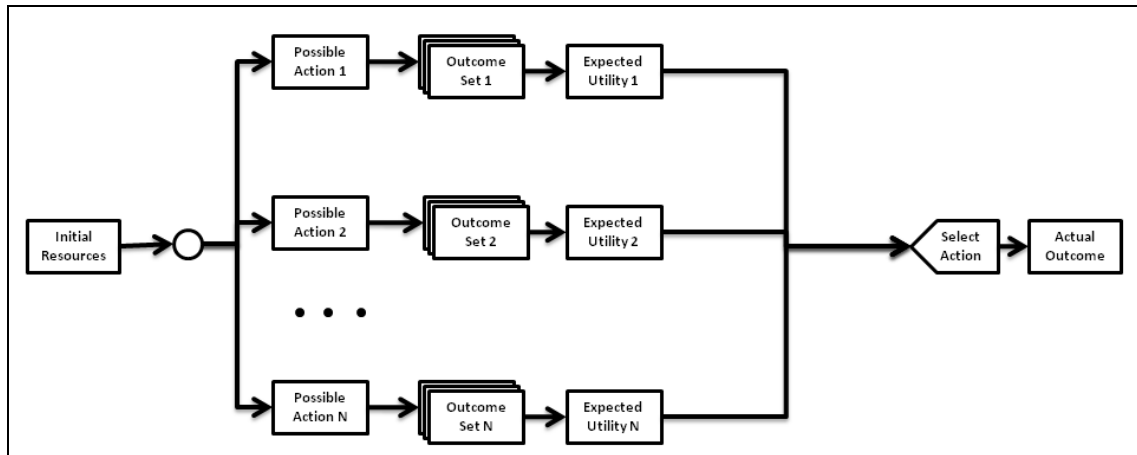
Figure 1: Classic Utility Model

The power of such a model in economic theory is undisputable. For example, using utility maximization as an axiom (combined with various assumptions regarding production functions) it becomes possible to demonstrate that a consistent set of demand and price values exist such that the result is general economic equilibrium (Allen, 1959). It is also possible to extend the model. For example, although not normally included in equilibrium models, it has been proposed that the independent values in the utility function (i.e., $x_1, x_2, \ldots, x_n$) need not be limited to goods and services but may also include less tangible elements such as power or the satisfaction of others (Simon, 1992). The idea of more abstract elements is also central to the notion of commodities in Becker's influential formulation of the household production function (Becker, 1976, p. 134). In addition, it has been widely recognized that human processing has significant limits and that available information is far less than the "perfect information" assumed in classical models. This recognition leads to a notion of bounded rationality (Simon, 1976) whereby a rational decision-maker limits utility maximizing activities once one or more satisfactory outcomes have been identified.

Because utility maximization is axiomatic to theories that postulate rational decision making, the underlying shape of individual utility functions has been researched extensively. Such research generally falls into one (or more) overlapping categories: the impact of risk on utility, the impact of time on utility, and the behavior of multi-attribute utility functions. Research into risk has focused on better understanding how people trade off costs and payoffs that include an element of randomness, either quantified (risk) or unspecified (uncertainty or ambiguity). The underlying theory involves a series of axioms (e.g., if outcome A is preferred to outcome B, then a gamble offering the probability of p for receiving A, 0 otherwise should always be preferred to a gamble with probability p of receiving B, 0 otherwise) that collectively define rational behavior under risk and ambiguity. In such choice situations, expected utility also becomes more critical than actual utility, as it is the former that is available prior to making the choice. One surprising result of this research has been the discovery of numerous anomalies that appear to call into question the rationality of actual human behavior (see Table 1).

A similar stream of research has examined utility tradeoffs over time, both with and without the presence of uncertainty. Corresponding axioms of rational behavior have been derived in the case of time (e.g., if payout A at t_0 is preferred to payout B at t_1, then payout A at t_n should be preferred to payout B at t_{n+1}). As was the case for risk and uncertainty, many temporal anomalies have also been identified (see Table 1). Hyperbolic discounting (Ainslie, 2001), which discounts

short-term delays far more drastically than the more economically elegant exponential discounting, has been proposed as an explanation for a number of these.

Table 1: Anomalies in utility models of rational decision making

Anomaly	Description	Example Reference
Preference Anomalies		
Framing effects	Individuals often express strong preferences when making choices between alternatives that are, in fact, identical.	Bell, Raiffa, & Tversky, 1988, p. 11.
Endowment effect	Upon acquiring an object, we value it far more than we did prior to acquiring it, leading to asymmetry between buying and selling prices.	Ariely, 2008, p. 130.
Anchoring	When establishing preferences, we typically anchor them to an available reference, even when that reference makes no sense (e.g., the last digits of a social security number).	Ariely, 2008, p. 28.
Preference reversal	When presented with two lotteries, individuals may prefer the less risky one yet be willing to pay more to participate in the riskier one.	Grether & Plott, 1979
Illusion of control	Individuals perceive that they exert far more control over random events than they actually do.	Langer, 1975, p. 231.
Cash effect	Experimental subjects react differently to cash rewards than to rewards easily convertible to cash.	Ariely, 2008, p. 220.
Availability bias	Judgments of likelihood are unduly influenced by the individual's ability to recall specific examples, which may or may not be representative.	Tversky & Kahneman, 1973, p. 163.
Expected value insensitivity	When presented choices of certainty and near certainty, we tend to prefer certainty even where payoff differences are substantial. As probabilities get very low, we focus on size of the payoff rather than its expected value.	Kahneman & Tversky, 1979, p. 267.
Temporal Anomalies		
Common difference	Identical payoffs separated by a fixed time period may change preference depending upon when the payoff is started. For example, you may prefer $3000 now to $4000 in a year, but also prefer $4000 in year 4 to $3000 in year 3.	Lowenstein & Prelec, 1992, p. 120.
Absolute magnitude effect	Ratios of preferences may change as payoffs change. For example, an individual may prefer $15 now to $60 in a year, but prefer $4000 in a year to $3000 now.	Lowenstein & Prelec, 1992, p. 121.
Gain-loss asymmetry	Losses are discounted at a greater rate than gains. For example, a study found an individual be indifferent to a gain of $10 now to $21 in a year, but be indifferent to a loss of $10 now and a loss of $15 in a year.	Lowenstein & Prelec, 1992, p. 122.
Delay-speedup asymmetry	Significant variations were found with respect to an individual's willingness to speed up consumption and delay consumption. In other words, they had to be paid more to move from t1 to t2 than they were willing to pay to move from t2 to t1.	Lowenstein & Prelec, 1992, p. 124.
Variance from future expectations	Individuals tend to consistently overestimate actual utility gains and losses when considering future events.	Gilbert, 2007, p. 102.
Sequence effect	Individuals generally exhibit a preference for a sequence of increasing payoffs that total to the same amount as decreasing payoffs.	Read, 2004, p. 435

Utility preference research investigating the impact of both risk and time generally focuses on a single attribute, money, as the underlying source of utility. Another stream of research, however, focuses on multi-attribute utility functions. The challenge presented by this research is that of ma-

thematical tractability. Generally, multi-attribute functions pose little problem provided that the utility from different attributes is linearly additive, i.e.,

(1) $U(x_1,x_2,\ldots,x_n) \equiv U_1(x_1) + U_2(x_2) + \ldots + U_n(x_n)$

Where interdependencies exist, however, the mathematics of utility maximization becomes much more complicated. For example, if an individual's utility from acquiring a bottle of foreign beer simultaneously depends upon the ability to acquire a bottle opener, then the utility of neither can be computed independently (particularly if no such opener is required for a domestic beer). One way of handling such problems mathematically is to include cross products in the utility function (Keeney, 1972). Unfortunately, as attribute interdependencies in the utility function grows, closed form solutions become impractical (Bell, 1979).

One approach that has been proposed to address the complexity of multi-attribute interdependency has been to use multi-attribute functions, referred to as objectives (Keeney, 1988, p. 471) or need-based commodities (Becker, 1976, p. 134), as arguments to the utility function. These functions, in turn, depend on actual resources. This approach allows us to construct a hierarchy of objectives that are independently summed within the utility function. The approach also forms the underlying basis of the analytical technique referred to as goal programming, a close cousin to linear programming, in which the object is to achieve:

(2) $\text{Max}_{x \,\varepsilon\, X}\, U(f(x))$

In this notation, x is a particular set of attributes from the set of all feasible attribute combinations X, $f(x)$ is a collection of objective functions $f_1(x)$, $f_2(x),\ldots,f_n(x)$, and U is a function that determines decision-maker utility (Dyer, 1972, p. 63). In contrast to utility research dealing with a single attribute, which has tended to focus of exploring how decision-maker preferences are impacted by risk and temporal proximity, much of body of multi-attribute utility research has been devoted to exploring the theory and mechanics of finding the state of maximum preference when a complex set of objectives is available to the decision-maker.

Motivation and Goal Setting Theory

When viewed in the context of the management field, expected utility is essentially a measure of a state's motivational potential. Stated another way, the underlying axiom of utility theory is that an individual will be motivated to make the decision (or set of decisions) yielding the state of maximum expected utility. Motivation, however, has been the subject of many investigations within the management literature—most of which bear little or no surface resemblance to neo-classical utility theory. Perhaps the most important of these research streams is goal setting theory (Locke, 2004, p. 124).

As its name suggests, the concept of a goal is central to goal setting theory. Goals possess a number of characteristics that have been found relevant in empirical research. These include:

- *Performance goals versus learning goals*. Performance goals relate to task outcomes. Learning goals (a.k.a. mastery goals) involve the development of individual skills.

- *Approach versus avoidance goals*. Approach goals are those where the individual strives to achieve a particular consequence. Avoidance goals involve attempting to prevent a particular consequence (e.g., a failing grade in a course).

- *Distal versus proximal goals*. Distal goals are desired outcomes, whereas proximal goals may be sub goals established as part of a plan to achieve a desired outcome.

- *Goal source:* Goals may be externally established, participatively established, or individually established.

- *Goal specificity.* The degree to which goal achievement is clearly defined. Within performance goals, subcategories include: a) ability goals, confirming or demonstrating the individual's skills, b) normative goals, where relative performance is targeted, such as coming in first place, and c) outcome goals, which involve more objective standards of performance (Grant & Dweck, 2003). Thus, a goal may consist of a specific target (e.g., sales of $150,000), a relative target (e.g., become a top 5 performer), or an open-ended target (e.g., sell as much as you can).

Goal setting theory posits that individual motivation can be explained and directed almost entirely through the goals that have been either internally or externally established. Unlike earlier motivational theories such as expectancy theory (Vroom, 1964)—which views motivation in terms of valence (how positive or negative an outcome is) and expectancy (the likelihood of an outcome)—goal setting theory proposes that the actual presence of a well defined goal, as opposed to some underlying psychological need, is critical to establishing motivation (Bandura, 1991, p. 264; Locke, 1978). In its simplest form, then, the relationship can be presented as follows:

(3)   State $\rightarrow$ Need $\rightarrow$ Goal $\rightarrow$ Motivation $\rightarrow$ Action

The resemblance of this model to the earlier utility-based decision model, presented in Figure 1, is obvious.

The goal setting literature is huge (e.g., over 100 different tasks, involving over 40,000 subjects, have been studied; Locke, 2004, p. 124) and has successfully demonstrated, in many contexts, how individual goals are instrumental in establishing motivation. Some important findings relating to the present paper are summarized in Table 2.

In considering the Table 2 findings, it is interesting to note a number of parallels between goal setting findings and some of the anomalies presented earlier in Table 1. For example:

- The impact of framing (i.e., how a choice is presented impacts decision-maker preferences) mirrors that of priming in goal-setting research.

- Significant differences between behaviors relating to gains versus losses.

- Both literatures observe a proximity effect in which closeness (e.g., in time, salience) weighs disproportionately on the individual.

- Both literatures find that low probability outcomes are very different from impossible outcomes in how they are perceived by decision-makers.

Another important parallel between goal setting research and utility theory involves the notion of limited attention. In utility theory, as previously noted, this manifests itself in the form of bounded rationality leading to constrained search behavior. In goal setting research, the implicit assumption is that multiple goals may be available and that commitment to a particular goal is critical to motivation (e.g., Klein, 1989, p. 886; Wofford, Goodwin, & Premack, 1992, p. 609). Within goal setting theory, this notion of goal intensity is also integral to a number of control theories of goal setting and motivation, which include both conscious and unconscious components (e.g., Johnson, Chang, & Lord, 2006).

The goal-setting perspective is particularly adaptable to the informing system model because goals represent a central component of tasks (Gill & Hicks, 2006). Thus, if we can relate utility to

goals, then we can also relate the role played by informing in the task completion system to utility.

Table 2: Important findings in goal-setting theory

Finding	Description	Example References
Priming	By unconsciously stimulating certain aspects of a goal, for example through having subjects read through a list of positive and negative words prior to participating in an experiment, substantially different goal performance and satisfaction may be realized.	Stajkovic, Locke, & Blair, 2006, p. 1175
Approach vs. Avoidance	Eagerness increases as you near completion of an approach goal, vigilance increases as you near completion of an avoidance goal.	Forster, Higgins, & Idson, 1998, p. 1129.
Specificity	Goals that are specific and difficult lead to better performance than a vague goal or no goal at all.	Latham & Locke, 2006, p. 332.
Mastery vs. Performance	Goal striving leads directly to enjoyment for mastery (learning) goals, and indirectly—through mental focus—for performance goals.	Lee, Sheldon, & Turban, 2003, p. 262
Commitment	Performance grows with goal commitment, especially for difficult tasks.	Klein, Wesson, Hollenbeck, & Alge, 1999, p. 886
Proximity	People exert more effort to achieve goals where progress has already been made, referred to as the "endowed progress effect".	Nunes & Drèze, 2006, p. 510
Difficulty	Increased goal difficulty leads to increased motivation and performance for approach goals, provided the goal is achievable, but less so for avoidance goals.	Janssen & Van Yperen, 2004, p. 377
Participation	Goals provide an important motivational regardless of whether or not the individual has participated in setting the goals. In many cases, participation doesn't appear to matter.	Latham & Steele, 1983, p. 416
Goals vs. Needs	Goals are better predictors of performance than personal achievement needs.	Bandura, 1991, p. 264

A Goal-Based Utility Decision Model

As described earlier, motivation can be characterized as the impetus to maximize expected utility. If one accepts the findings of goal setting research it therefore stands to reason that a utility function could be constructed using goals, rather than other attributes (such as money), as its arguments. The validity of such a function would depend on several axiomatic assumptions, namely:

1. At any given time, an individual holds a set of conscious goals that can be articulated. Such goals may derive from many sources, including individual needs and external sources, such as the well-being of friends and family, the goals of organizations and communities with which the individual is affiliated, and even broader sources, such as spiritual goals, national goals, and environmental goals.

2. Utility derives from two processes: the direct satisfaction of needs and progress in the pursuit of the individual's goals, both in absolute terms and relative to expectations. For the purpose of the model, needs are treated as concrete biological goals (Politser, 2008, p. 30), allowing all utility to be treated as arising from goal-based sources.

3. Attending to all goals concurrently would exceed the individual's processing capacity by many orders of magnitude. As a consequence, mechanisms for prioritizing, selecting, scheduling, and refining active goals are critical elements of reasoned behavior. These come in two forms: unconscious processes that establish preferences and heuristics—both specialized and general-purpose—that tend to be consciously applied. We refer to the latter as meta-reasoning behaviors.

4. As a result of axiom (2), we may assume that the decision to engage in explicit meta-reasoning behaviors is also goal driven. Goals of this type will be referred to as meta-goals.

Of these axioms, the second needs the most justification. Through introspection, most of us will probably concede that satisfying certain biological needs—such as hunger—often provides a source of utility. In nearly all animal species, that may well be the end of the story. One of the key differences between human and animal cognition, however, is the ability to think about the future (Gilbert, 2007, p. 4). To do so, we have developed specialized brain areas, such as the frontal lobe, whose principal function appears to be planning (Gilbert, 2007, p. 13). Furthermore, biological evolution tends to take such newly developed capabilities and superimpose them on top of existing capabilities, meaning that both planned and unplanned actions exist in parallel. In fact, it took considerable time for scientists to discover that a damaged or disconnected frontal lobe—such as occurs when a patient has a lobotomy—typically results in inability to plan; patients otherwise seemed quite normal (Gilbert, 2007, p. 11).

In addition to providing the capacity to plan, however, evolution needed to evolve a corresponding drive to use that capacity. One consequence of this seems to be that the frontal lobe can also be a major source of anxiety (which is why lobotomies were performed, since cutting the connection to the frontal lobe produced immediate declines in patient anxiety unobtainable through other means; Gilbert, 2007, p. 13). What the goal-based utility model therefore proposes—based upon both physiological evidence and the extensive goal-setting literature—is that goals provide an organizing framework around which most planning takes place and that utility, both in the form of reduction of anxiety and positive satisfaction, is derived from pursuing such goals.

The actual goals driving the goal-based model would result from a number of sources. Goals are created as responses to underlying needs, consistent with the goal-setting model. Examples of such needs include the drive to acquire, to bond, to learn, and to defend (Lawrence & Nohria, 2002). Even lower level need-based goals—such as those towards the bottom of Maslow's hierarchy—would routinely arise from fundamental biological drives (e.g., hunger, sex, security, social interaction) as a consequence of planning for future satisfaction of those needs. Where planning is not required, a direct path from need satisfaction to utility can occur. The model would predict, however, that the individual's utility gain from directly satisfying a need versus progressing towards a conscious goal derived from that same need could be quite different. This is consistent with findings that reasoning about a goal changes its utility (e.g., Dijksterhuis, Bos, Nordgren, & van Baaren, 2006).

Higher levels of the need hierarchy, much more uniquely human, would drive the formation of other goals. Fundamental desires, such as those for power, independence, acceptance, and status (Reiss, 2000), would drive the formation of achievement goals, such as career advancement. Other desires, such as curiosity, order, and tranquility (Reiss, 2000), could lead to a drive to reduce uncertainty. These higher level goals would tend to be conscious in form.

Meta-reasoning processes, invoked by the drive to reduce uncertainty, would focus on goal identification and planning-related activities such as prioritization and scheduling. For example, the environment may present the individual with an unexpected opportunity stimulus. In some cases,

a special-purpose script—conscious or automatic, acquired through past experience or education—for translating that stimulus into a goal and creating an action plan may exist. Because such an opportunity can come in many shapes and forms, however, each decision maker must also possess a set of general purpose heuristics for recasting novel stimuli into utility preferences and action plans. As previously noted, general-purpose preference heuristics have been the subject of considerable study (e.g., Kahneman, Slovic, & Tversky, 1982; Payne, Bettman, & Johnson, 1993). Where preference determination is required, unconscious and conscious processes may both be involved. In fact, it has even been shown that for certain types of choice problems relying on unconscious processes may lead to better results than applying available heuristics (Dijksterhuis et al., 2006; Gladwell, 2005; Wilson & Schooler, 1991). Planning heuristics, on the other hand, would nearly always be conscious. These would include general purpose strategies such as decomposing a complex goal into subgoals (e.g., Simon, 1981) and various search heuristics (e.g., Newell, 1990).

Particularly critical to the goals-based model is the notion that goals can also be adopted from other individuals and communities with whom the decision maker interacts. Simon (1992) argues, for example, that such adoption of communal goals is critical in explaining the formation of organizations and altruism. It would also be nearly impossible to explain behaviors such as enthusiasm for a particular sporting team, religious tithing, or nationalism without the notion of communal goals. Simon quotes Gary Becker as saying:

> I decide whether to turn off my reading lamp in bed by comparing the utility of my pleasure in reading with the amount of utility I receive from my wife's being able to sleep. (Simon, 1992; p. 74)

Experimentally-based psychological research seems to confirm that altruistic goals are a source of motivation separate from egoistic goals (Batson & Shaw, 1991). There are also evolutionary arguments that can be made that altruistic behavior could, ultimately, enhance individual fitness (Gandolfi, Gandolfi, & Barash, 2002). Because cooperative behaviors can increase efficiency (e.g., time spent defending your land against competitors is time that cannot be spent farming) and short-term altruistic behaviors (e.g., helping a neighboring family defend their land against competitors) would tend to promote such cooperation over the long term, drives that lead to the formation of altruistic behaviors could form an evolutionarily stable strategy (ESS). This does not necessarily imply that the "altruism gene" would be universally present throughout the population. Rather, it means altruistic individuals (competing against entirely selfish individuals, such as sociopaths) would exhibit sufficient fitness to remain a percentage of the overall population indefinitely.

The notion that we, as human beings, might have an underlying biological drive leading us to derive pleasure from participating in and observing progress towards the goals of others is both intuitive—how could we enjoy reading a novel or watching a sporting event if such a mechanism were not present?—and profoundly subversive, potentially undermining important elements of the theoretical justifications for both transaction cost economics (e.g., Moschandreas, 1997) and agency theory (e.g., Simon, 1991). There also appears to be a physiological basis for believing how such an affinity could exist. Within the brain, mirror neurons respond to observing the acts of others in much the same manner as they do when we perform the acts ourselves (Shermer, 2008, p. 132). Thus, it would make sense to speculate that their presence could also lead to utility as we observe external goals being achieved.

A goal-based utility function would differ in purpose from other objective-based utility functions (e.g., Dyer, 1972; Keeney, 1988) mainly in that its principal objective would be to describe and predict individual behaviors, rather than to establish an approach whereby utility could be maximized. It differs from neoclassical utility functions—which are not precluded from incorporating

goals as attributes (Simon, 1992)—in that it proposes that an understanding of individual goals and meta-goals is a prerequisite to meaningful explanations and predictions of preferences and decisions. Conceptually, this difference can be represented by adding a stage that specifically maps outcome sets to the goals that they serve as part of the utility maximization process. For example, if we assume that a decision-maker is pursuing M different goals under conditions of uncertainty, we might represent this as illustrated in Figure 2.

A critical difference between the goal-based model and the neoclassical model stems from the concept of goal allocation. The idea here is that a given set of outcomes may map to different utilities depending upon how the changes are allocated among the decision maker's various goals. For example, if a decision maker's gas tank is empty, the first step in going on any road trip would be to stop at the gas station. What the goal-based utility model would predict is that the "utility" associated with choosing the action of going to the gas station will depend heavily on what goal or goals (e.g., what destination, what anticipated activities are associated with the destination) the decision-maker intends to pursue upon completing the gas filling action. Naturally, the mapping between actions and utility is one-to-one if a given action can only support a single goal. If, on the other hand, a particular action (e.g., getting an undergraduate degree) supports multiple longer-term goals that are mutually exclusive (e.g., getting a job, going on to medical school, going on to law school), then the utility associated with making progress towards the degree may depend heavily on the relative priority associated with each of the longer-term goals. As priorities change, so would utility. Furthermore, the goal-based model suggests that framing "anomalies" in which decision maker's express inconsistent preferences for the same resource-based outcome are better explained in terms of differing goal allocations being in force rather than as demonstrations of decision maker irrationality.

The model in Figure 2, as presented in its purest form, might easily be described as "pegging the meter" of cognitive implausibility. The number of possible goal allocations is, effectively, infinite if we allow allocations where each goal is given a percentage priority. Even if we limited ourselves to rank ordering goals, if there are M goals there are M! (M factorial) allocation combinations. To reduce the choice process to cognitively manageable proportions, we therefore need to postulate mechanisms that drastically reduce the number of alternatives being considered. Thus, the model postulates that at any given time we are attending to a relatively small number of goals. These goals are determined by existing plans and environmental stimuli. The proposed impact of attention is to:

1) Suppress the number of actions we consider, leaving only those relevant to goals we are attending.
2) Suppress the number of allocations that we attend to, focusing only on those with attended goals
3) Suppress the number of allocations we consider for each outcome set, ignoring those where the goal-outcome connection is expected to be weak.

In Figure 2, these are shown as inhibiting links (dotted lines ending in a circle) from the attention box to the circles where alternatives and combinations are generated.

In the most extreme case, the inhibitory effect would lead to focusing on a single goal, devising a sequence of actions (i.e., a plan) towards achieving that goal and experiencing utility directly proportional to the progress towards the achievement of that goal (and nothing else) as the sequence of activities continues. In such a circumstance, the utility model could—in effect—converge to the economic utility model, as progress towards the goal might be a function of external values (e.g., how many papers do I have left to grade?).
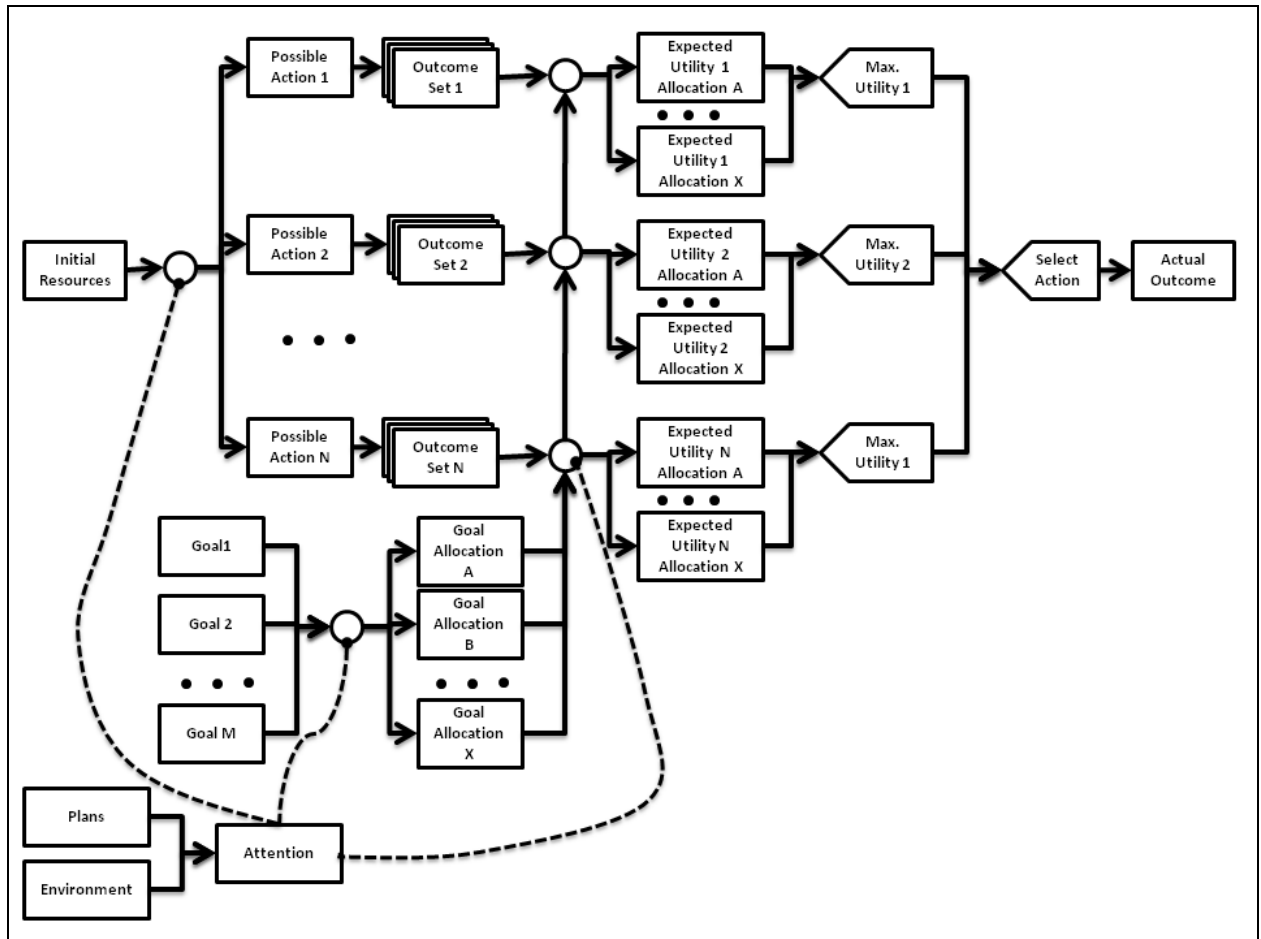
Figure 2: Goal-Based Utility Decision Model

In a more realistic scenario, the decision maker would have multiple goals but these would differ in the degree to which they are active. A large percentage of actions would be a consequence of goal-directed plans, but periodically a more complete reassessment of available actions would be made in response to conditions such as variances between actual and expected outcomes within a plan, plan completion, need-based signals activating other goals (e.g., hunger leading to the goal of going out to eat), or external stimuli activating other goals (e.g., while shopping for one item you notice an item you also wanted on sale). The decision to engage in plan reassessment can be referred to as activating a meta-goal and, according to the model, our choice to engage in such planning activities would only occur when we perceive that doing so will lead to a successful plan (providing utility by moving us towards the meta-goal) that will improve our progress towards other goals or, possibly, where the act of planning serves to reduce our anxiety regarding how long it has been since we last planned. The price we pay for this form of bounded rationality is that the utility experienced as a consequence of our actions will depend heavily on the specific goals we are attending, meaning that the same resource state could lead to different utilities based upon what goals are being pursued: a classic scenario for framing.
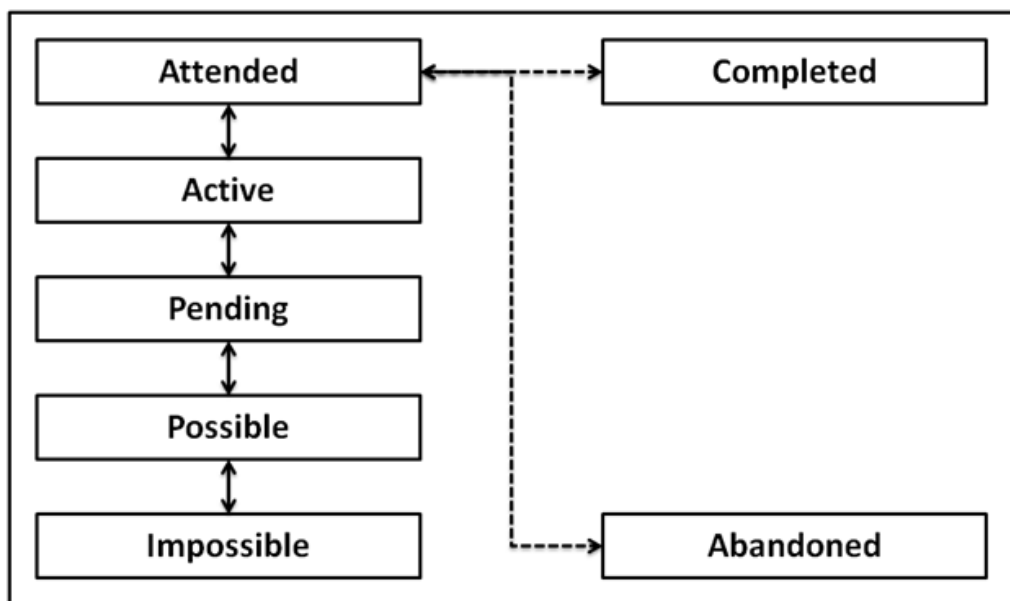
In considering how the attention and goal allocation elements that contribute to the utility function might be constructed, goal setting research makes it abundantly clear that we would need to consider:

- The individual's commitment to the goal (e.g., Klein et al., 1999)

- The degree to which the individual is consciously attending to the goal. For example, goal priming has repeatedly been shown to influence preferences and decisions in many experimental settings (e.g., Stajkovic et al., 2006)
- The likelihood of achieving the goal, which may be a function of both perceived goal difficulty (e.g., Janssen & Van Yperen, 2004) and considerations of external uncertainty.

As a preliminary model, based on the existing literature, the attention signal feeding into the utility function might—within its arguments—partition individual goals into categories, shown in Figure 3. Of the five categories stacked vertically, the highest category, attended goals, would consist of the one (or, possibly, few) goals actively being processed by the individual at a given time (i.e., in working memory). The next category, active goals, consists of goals that the individual is actively pursuing but which are not currently attended. For example, a pre-med undergraduate student may have getting into medical school as an active goal without concentrating on it every waking moment. The middle category, pending goals, represents those goals to which the individual has significant commitment but towards which no action has yet been taken. An individual might, for example, have a goal of eating at some specific restaurant the next time he or she visits New York City yet not engage in any action towards fulfilling that goal until the visit actually occurs or is planned. The fourth category, possible goals, represents those goals to which the individual has no firm commitment but to which the individual has some attachment. Such a category is necessary in order to incorporate the construct of "hope", which has been observed to have powerful psychological impact (e.g., Snyder, 2002). The final category on the stack, impossible goals, represents the collection of goals that might have served to motivate the individual but that have been deemed infeasible.

Figure 3: Goal Attention Levels

The remaining two categories, completed goals and abandoned goals, would hold goals no longer active, either by virtue of completion or by virtue of failure to achieve or abandonment. In both cases, we would expect a residual effect on utility that would rapidly decay with time, since utility effects of success and failure to achieve goals tend to decline quickly—indeed, more quickly than we expect (Gilbert, 2007, p. 102). We may, however, choose at some point to revisit these goals mentally and re-experience some of the same utility—to relive our success or, perhaps, to learn from our failures (Elster & Loewenstein, 1992). Doing so appears to be possible because

much of the sensory and emotional circuitry the brain uses when we experience something is also used when we recall the same experience (Gilbert, 2007, p. 120).

Learning Architecture of the Utility Function

The previous section focused on how attention impacts utility in the context of an overall decision-making model. In this section, we consider more closely the type of cognitive architecture required to implement the utility function itself. In the goal-based utility model, utility would be a function of at least three distinct elements:

1. The attractiveness of the goal outcome (roughly equivalent to the valence in expectancy theory), which in turn would be influenced by goal type (e.g., learning vs. performance, approach vs. avoidance goals, specific vs. general goals will tend to have different profiles).

2. The individual's perception that he or she is making progress towards the goal (including activities that create a more concrete path towards the goal, such as planning). Where the individuals own actions are seen as being instrumental in achieving such progress, the utility would be even greater, as the desire to exercise control is another example of a fundamental human need (Gilbert, 2007, p. 20).

3. The degree of commitment to the goal, as approximated by its goal attention level (Figure 3), with goals moving up (towards attended) increasing utility and moving down (towards impossible) reducing utility. This is consistent with the observed psychological phenomenon of preference increasing as choices become increasingly irreversible (Frey, Kumpf, Irle, & Gniech, 1984; Gilbert, 2007, p. 183-184). This effect could also lead to the observed phenomenon of increased effort being applied towards goals that are approaching completion (e.g., Nunes & Drèze, 2006).

As mentioned in the introduction, a major weakness in any utility model is the failure to incorporate a mechanism for learning. Specifically, as humans we have the ability to establish preferences in both familiar and unfamiliar settings. Any plausible utility model therefore needs to establish a mechanism capable of handling both types of preferences.

To accommodate learning, we can view the utility function as existing on three levels, depending upon the familiarity of the situation. At the lowest level—that level where the context is so familiar that utility-based choices are unnecessary or can be made without conscious effort (e.g., automated activities; Shiffrin & Dumais, 1981) —the function might look something like that proposed by economists, with arguments consisting of a set of values and little need to consciously attend to goals. In this condition, utility becomes a type of stimulus-response phenomenon where it is unlikely that the decision maker will be able to justify his or her utility. The pleasure derived from engaging in a practiced skill (e.g., playing the piano) might be an example of utility at this level. In the previous section, this was described as the single-goal focus.

The intermediate level, which would be particularly consistent with the goal-based decision-making model described in the previous section, would have arguments consisting of specialized goals, with utility deriving from activation levels and goal progress. At this level, decision makers should be able to offer an explanation of their decision-making process, since goals are conscious in nature. This would also be the domain in which the findings of the goal setting literature would be most applicable.

At the highest level, we find ourselves encountering unfamiliar situations. For this level, there is a particular need for general-purpose mechanisms for estimating utility. Such a mechanism can be constructed using detectors—cognitive units that identify choice attributes that are expected to

contribute to or detract from utility—and accumulators, which consolidate signals from detectors and other accumulators in order to assess the overall fitness of a particular choice.

There is a long history of cognitive models—originally focused on word identification—that have employed the detector-accumulator model. The original architecture proposed was Pandemonium (Selfridge, 1958), which treated cognitive architecture as a series of layers. At the bottom layer, feature detectors (referred to as daemons) identified letter features—e.g., slanted lines, vertical lines, horizontal lines, and arcs—that sent signals to a second layer containing accumulators associated with each letter that, in turn, sent signals to word accumulators that identified words. This original model was then extended in HEARSAY and then in the interactive activation model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982). More recently, the leaky accumulator model (Usher & McClelland, 2001, 2004) has been specifically applied to explaining precisely the types of inconsistency of choice problems that were identified in Table 1.

Because our interest here is in understanding how utility behaves in general, we'll focus on ways in which these models are similar, rather than on how they are different. In common, all the models assume:
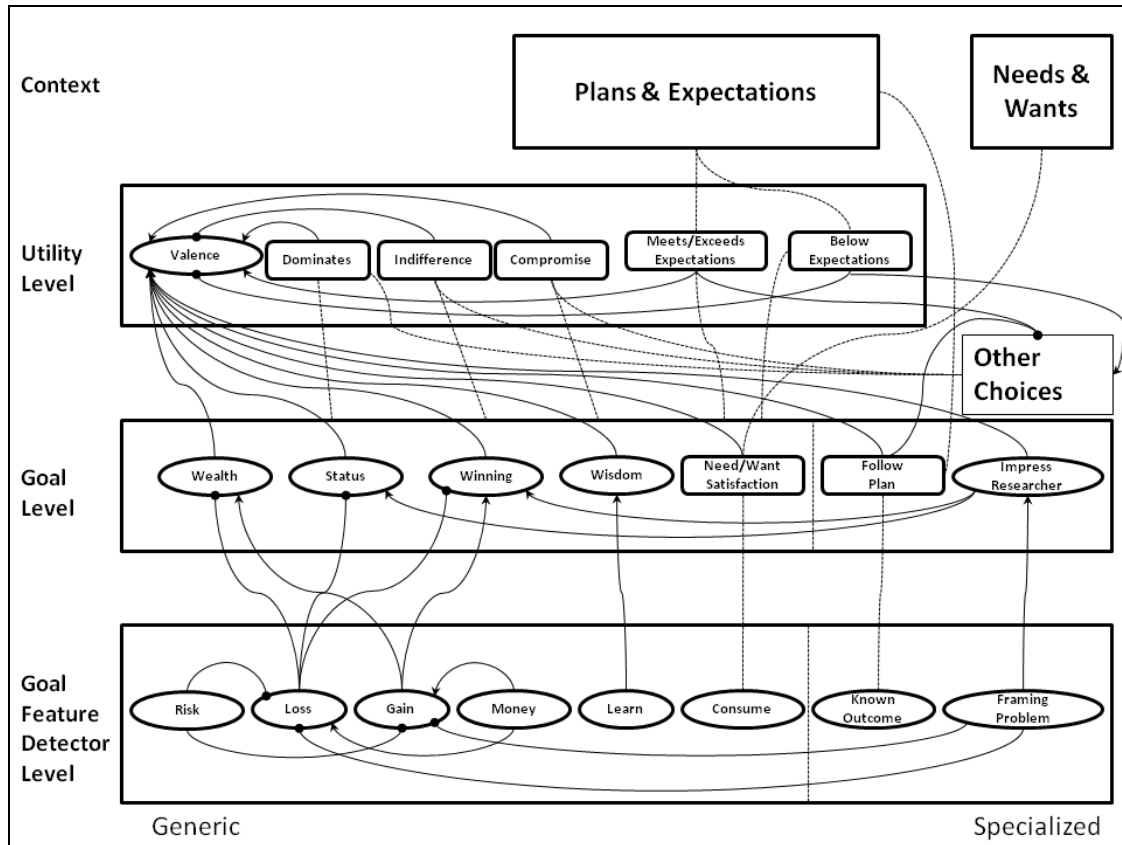
- Parallel processing of inputs.
- The existence of feature detectors and the ability to develop new feature detectors.
- The existence of layers of elements (e.g., detectors, accumulators) at different levels of abstraction.

Common to the later models, heavily influenced by findings in neural networks, are the following additional similarities:

- A substantial amount of "knowledge" is contained in the strength of linkages between system elements.
- Linkages can both excite and inhibit accumulators and detectors
- Learning occurs through changes in the strength of connections between elements.

Where models tend to diverge is in relation to the process by which links change, the degree to which symbolic versus connectionist processes dominate, and the nature of activation signals (e.g., linear versus non-linear). For the present purposes, however, we need not dwell on these differences; ultimately these will most likely be settled in the domain of neuroscience.

Our overall utility model for unfamiliar choices is illustrated in Figure 4. At the bottom layer, we have a set of generic goal feature detectors that can be applied to nearly any choice problem as well as a few examples of problem specific goal feature detectors. These detectors would be applied to a specific choice to identify attributes relevant to satisfying goals. The next level has a collection of generic goals—which is to say goals that are always waiting to be achieved—plus specific goals, with a few examples provided. The most important source of activation for nodes on this level is signals from the feature detectors of the lower level. Thus, the arrow going from the *Gain* feature detector to the *Wealth* goal node signifies that when a gain is detected in the particular choice, activity in the *Wealth* node increases. On the other hand, the inhibiting connection (ending with a circle instead of an arrowhead) between *Loss* and *Wealth* signifies that a detected loss will decrease activity in the *Wealth* node. The next level is the utility level, which takes signals from the various goals activated at the goal level (signified by the cluster of arrows on the left of the figure going into the *Valence* node) and sums them to produce a valence whose strength signifies utility. Finally, the top level includes factors that may ultimately impact utility, such as plans and visceral factors (e.g., decision-maker needs such as hunger, which we would expect should increase the utility of any options that lead to food).

Figure 4: Multi-level utility model emphasizing unfamiliar problems

Consistent with activation models, such as the interactive activation model (McClelland & Rumelhart, 1981) and the leaky accumulator model (Usher & McClelland, 2004), there are a number of aspects of the model that extend beyond the basic movement of activation upwards from layer to layer. At the bottom layer, for example, the *Risk* detector sends inhibiting signals to both the *Loss* and *Gain* detectors. This means that as risk is detected, the strength of both gain and loss signals will be reduced, since they lack certainty. Some nodes are also constructed to be comparator nodes, shown as rounded rectangles. These nodes take connections, shown as dotted lines, from various sources and produce an activation or inhibition signal based on some comparison. For example, the *Need/Want Satisfaction* comparator node on the goal level takes consumption (or expected consumption) signals from the detector level—simplified to appear as a single node, although many types of consumption could obviously be involved—and compares these to signals from the want/need area (also simplified as a single line). Activation signals are then produced reflecting the degree to which consumption (or other activity) associated with the choice option being considered matches the current need profile.

Particularly strong support for a model such as the one proposed is provided by framing experiments. The framing phenomenon occurs whenever the choice between identical outcomes is influenced by how the alternatives are presented. The same phenomenon extends to situations where the presence of irrelevant or "decoy" alternatives exert a significant impact on our choices. Framing is particularly vexing to traditional economic utility models because it violates the mapping between outcome and utility.

A typical framing situation—and there are numerous examples of these—is presented in the form of two sets of choices that lead to the same outcome and but are presented in different ways. Con-

sider, for example, the following (adapted directly from Tversky & Kahneman, 1988, p. 173-174):

First choice problem:

Assume yourself richer by $300 than you are today. You have to choose between:

> A. A sure gain of $100
> B. 50% chance to gain $200 and 50% to gain nothing

Second choice problem:

Assume yourself richer by $500 than you are today. You have to choose between:

> A. A sure loss of $100
> B. 50% chance to lose nothing, 50% to lose $200

The actual results of this experiment were that most (72%) chose A in the first problem, while most (64%) chose B in the second. This presents a bit of a problem for any utility model based on values, since—in fact—options A and B in both choice problems lead to precisely the same outcome: a $400 gain in choice A, a 50-50 distribution between $300 and $500 in the case of choice B.

Now, let's consider this experiment under the architecture that was illustrated in Figure 2. First, however, we need to make some assumptions. We begin by assuming that we evolved as a fundamentally peaceful species that needs to engage in three basic activities to survive: acquiring resources, consuming resources, and defending resources from threats (which might include predators, less peaceably evolved humans, and more mundane threats such as spoilage and pests). Not coincidentally, these resemble the capabilities viewed as necessary for the general-purpose agent: offense, defense, and a reservoir (Holland, 1995, p. 101). Supporting this conjecture is ample psychological evidence for an asymmetry between gains and losses, incorporated into prospect theory (Kahneman & Tversky, 1979) and SP/A theory (Lopes & Oden, 1999) and supported by some brain studies (Polister, 2008, p. 81). In fact, observed reactions to loss and gain are fundamentally different (Gilbert, 2007, p. 145), as are our attitudes towards what we already have and don't have (the "endowment effect"; Ariely, 2008, p. 129). Thus, there is no reason to expect that activation signals from *Gain* and *Loss* detectors will be symmetrical. We also expect the *Risk* detector node, which acts to suppress the activation of the remaining frames in the event that uncertainty is detected, will generate inhibition signals varying roughly with probability but less finely tuned (Slovic, Finucane, Peters, & Macgregor, 2006).

Our second assumption is that we have no specialized goal frames for evaluating gambles (gambling being unusual in that it is an activity where acquisition and loss of resources occur through the same process) and, most importantly, that we have no specialized avoidance goals that relate to being embarrassed by the types of inconsistencies produced by framing experiments. We will later consider how such a goal, already shown as a "specialized" node in Figure 4, might impact utility.

Under these circumstances, when we consider our first choice problem, we'd expect the following:

- Option A: *Gain* node sends out strong signals. All other detectors keep quiet.
- Option B: *Gain* node sends out signals that are suppressed by the *Risk* node.

Based on this model, activation would likely be higher for Option A in the event that risk adversity exists, consistent with the actual findings. What this would imply is that the *Risk* node sends

out inhibitory signals stronger than justified by probabilities alone. Now, let's examine the second pair of choices:

- Option A: *Loss* node sends strong inhibiting signals. *Gain* node sends out a positive signal. All other goal frames keep quiet.
- Option B: *Loss* node sends signals that are inhibited by the *Risk* node. *Gain* node also sends out the same positive signal.

If the inhibiting impact of the *Risk* node has the same inhibitory strength on *Loss* node as it did on *Gain* node, the second option is likely to yield the stronger activation. If strengths differ, a different pattern of responses may emerge (which is also consistent with the fact that not all subjects chose the dominant pair of options).

Naturally, all this changes once the experimental subject has specialized knowledge of framing experiments. Under these circumstances, it is very likely that the subject will recognize the familiar pattern common to such experiments. This specialized *Framing Problem* detector will then activate the *Impress Researcher* specialized goal node and send inhibitory signals to the *Gain* and *Loss* nodes so that they do not confuse the subject—as they were originally intended to do by experimental design. As a consequence, the subject will therefore make the same choice in each pair (unless some other specialized goal node exists, such as a "rebellion goal" that activates based on detecting an opportunity to disrupt the experiment).

A particularly pervasive category of framing occurs as a consequence of detecting favorable comparisons and similar options among choices. In the case of favorable comparisons, when three options are presented to a decision-maker and two are similar, with one clearly dominating the other, decision-maker preference tends to skew towards the dominant option in the pair, regardless of the characteristics of the third. This is referred to as the attraction effect (Usher & McClelland, 2004). In the case where three options are presented and two are similar (with neither dominating), decision-maker preference skews towards the third, non-similar option. This is known as the similarity effect (Usher & McClelland, 2004). In the case where one option could be viewed as a compromise between the other two, that option is disproportionately selected. This is known as the compromise effect (Usher & McClelland, 2004). Interestingly, all three of these observed behaviors represent reasonably sensible approaches to highly unstructured choice making. The attraction and compromise heuristics tend to assure that the option chosen will be better than at least one of the other two—improving the decision maker's odds. The similarity heuristic, on the other hand, brings decision-making to a close more rapidly.

In Figure 4, these three phenomena are handled by generic comparators at the utility level. The detection of option dominance or of an option serving as a compromise are both proposed to increase utility activation (the *Valence* node). Where an option is found to be indifferent to other options—the similarity effect—utility is inhibited.

It should be noted that Figure 4 is provided to illustrate the architectural structure only and is by no means intended to be complete. In fact, a number of other general-purpose goal nodes and goal feature detectors can be postulated based on research conducted in a variety of contexts. For example, we might postulate detectors that feed into an "Engaging Activity" goal node that include:

- Identify situations where we gain control (+) and lose control (-) as a consequence of a decision, since control has been found to be a powerful motivator (Gilbert, 2007; Gill, 1996).

- Identify situations where the expected consequences of a decision lead to cognitive arousal that is likely to be too low (-), too high (-) or suitable for optimal engagement (+), with arousal being another general motivator (Gill, 1996).

- Identify variety in associated rewards and activities (+) versus uniformity (-), supported by the job enrichment literature (Hackman & Oldham, 1980) and as demonstrated by finding related to individual preferences for rewards (e.g., Gilbert, 2007, p. 129).

We could also postulate an altruistic goal node (e.g., Batson & Shaw, 1991) that would have associated feature detectors. Another example might be a detector that reacts to the particular presence of money, consistent with a large number of examples by Ariely (2008), who found that its mere mention was sufficient to "cause market norms to emerge" (p. 74). This might be an example of a goal frame detector that amplifies the signals from other detectors (e.g., gains, losses).

For the proposed architecture to be plausible, as specialized learning progresses, the generic model should exhibit convergence with our intermediate (specialized goal-driven) and low level (stimulus-response) utility models. Such convergence is envisioned happening as follows. Specialized goals and goal detectors that are acquired through experience would inhibit irrelevant or distracting general detectors and goal nodes, increase the activation of those that are relevant, and send signals directly to the valence. This would lead to utility deriving almost entirely from specialized goals. In addition, specialized comparator nodes would examine goal features and task state with respect to expectations associated with existing plans. As shown in Figure 4, these could act at multiple levels to increase activation where activities (detector level and goal level) were consistent with plan and where choices and outcomes were consistent with plans. They could also inhibit other choices where performance met or exceeded the active plan. Where inconsistency with plans was encountered, on the other hand, inhibiting signals could be sent to the *Valence* node and activation signals to other task choice valences could also be sent. This process would tend to lead to a change in selected action. Initially, such signals would reward goal progress in combination with other goals. Eventually, for highly routine tasks, these signals might suppress all other goal-related signals and drive utility for the entire process, leading to the predicted low level stimulus-response behavior.

It must be emphasized, once again, that Figure 4 is for illustrative purposes only. In a truly realistic neural model, connections, of varying strengths, would tend be much more dense between nodes and the nodes would be considerably less arbitrary in their names and functions. What the illustration does demonstrate, however, is that a multi-level architecture along the lines described is conceptually feasible, accommodates learning, and can explain many phenomena that are viewed as anomalies in traditional neoclassical utility models. These are the key elements of plausibility that need to be established for our purposes. The details of the example model—which are certainly oversimplified and likely to be wrong in many specific ways—would need to be verified and modified over the course of future research projects.

Divergences between Goal-Based and Neoclassical Utility Functions

Since adding explicit goals to a utility model increases its complexity when compared with neoclassical models, it would be hard to justify adopting a goal-based approach if the predictions of the two models did not differ in some non-trivial ways. For example, in many ways, the goal arguments of the proposed utility function are similar to Becker's (1976) commodities—something particularly true of the generic goal nodes proposed for the unfamiliar choice model. These commodities, however, are presumed to be need-based and relatively stable over time (Gandolfi et al., 2002). In contrast, the goal-based model anticipates a continually changing internal landscape of goals and considerable attention-driven variation in utility. Thus, while neoclassical models may be more appropriate for understanding aggregated macroeconomic phenomena, the goal-based utility model is distinctly focused on individual behavior, such as understanding and predicting preferences.

With respect to the framing behaviors, differences between the two models are immediately apparent. In the neoclassical model, the existence of a difference in utilities between two identical payoffs is a serious violation of the axioms and, therefore, can be viewed as irrational. In the goal-based model, on the other hand, such differences are expected, as would be preference reversal (e.g., Grether & Plott, 1979). Interestingly, whereas neoclassical models may concern themselves with the need to train decision-makers to reframe problems (e.g., Somon, 2004, p. 396), in the goal-based utility model the source of framing is viewed as a mechanism for coping with situational lack of structure. As familiarity with a situation increases, the mechanisms that lead to framing would be suppressed by more specific goals. In the informing system context, this would suggest that systems built around repetitive tasks might be less susceptible to presentation-related biases than might be expected based upon experimental results.

Another area where goal-based utility would differ from neoclassical models is with respect to time. As noted earlier in Table 1, neoclassical models have a great deal of trouble constructing realistic time preferences as a result of immediacy effects, inconsistent discounting, and variations in time preference for different magnitudes of rewards. In the goal-based model, we would expect a number of effects. For modest monetary reward amounts, we would likely have an intuitive—perhaps even unconscious—sense of how the money could be employed in achieving our currently active goals (i.e., what we'd spend it on). This would make immediate money particularly "valuable" in utility terms. That effect would wear off quickly with delays in payout, however, resulting in a pronounced immediacy effect. For much larger sums, we would generally expect the immediacy effect to be much less pronounced, since a large sum would normally require significant rethinking of goals in both the immediate and delayed payment cases. (As an exception to this, the goal-based model would also predict that the immediacy effect would continue to be pronounced under circumstances where the decision maker happened to have an attended or active goal that could efficiently benefit from a large infusion of cash). A further prediction of the goal-based model would be that immediacy effects would be less pronounced for delays in non-cash payouts, such as restaurant gift cards, since such a payout identifies the goal to be accomplished (quelling hunger) in both near and long term. (Again, the predicted exception would be in the case where the decision maker was hungry at the time the choice was presented.) In fact, we might speculate that it is the very fact that such gift cards encourage recipients to attend to goals that are more specific than money that accounts for their popularity with givers—as it is hard to understand why instruments that limit recipient flexibility would be given in preference to cash under any neoclassical utility model.

The relative importance of discounting also diverges between the neoclassical and goal-based utility models. In the neoclassical model, incorporating the time dimension requires the estimation of utilities at many future points in time and then incorporating these into the decision making framework. While the goal-based model does not necessarily preclude such elaborate calculations to establish time preference, it does not necessarily depend upon them. Although goals can be achieved either in the near or distant future, progress made towards such goals would contribute towards utility in the present. Savings, then, would not necessarily be treated as a form of deferred gratification (sacrificing current utility for future utility). Rather, the goal-based model would assert that the present utility gained from making progress towards the savings goal is greater than the utility that would be derived from current consumption. This notion is consistent with social self-regulation theory, which assumes "behavior is directed by cognized goals, not pulled by an unrealized future state" (Bandura, 1991, p. 248). There is also some biological evidence that delayed and uncertain outcomes employ a great deal of common circuitry in the brain (Politser, 2008, p. 54) and that our perceptions of differences in distant time are fuzzy (Gilbert, 2007, p. 107), much fuzzier than would be implied by the computations entailed in neoclassical discounting models.

Another important difference between goal-based utility and neoclassical utility involves the presumed impact of rising income. In neoclassical economics, the problematic "Easterlin Paradox" involves the observation that, across industrialized nations, overall happiness—which closely corresponds to utility—does not improve with national income, although, within a particular country, happiness generally rises with income (Clark, Frijters, & Shields, 2008). Similarly, average national happiness does not necessarily rise as per capita income rises. This is problematic for any resource/consumption based utility function. The usual economic explanation involves postulating that it is relative income, both with respect to a set of peer individuals and with respect to one's own past performance, that actually produces utility (once subsistence needs have been met). Two comments can be made with respect to this paradox and its economic explanation. First, a goal-based utility model would predict minimal utility effects from rising income except to the extent that it permits individual goals—which could, of course, be dominated by wealth accumulation or consumption goals but which are likely to be much broader in practice—to be better achieved. The obvious exception here would be a predicted boost in utility as individual income rises just above subsistence, which would be the point at which many discretionary goals would start to become possible, an economic phenomenon that has been observed (Clark et al., 2008).

With respect to utility gains with rising income within a particular nation, the goal-based model would take a different perspective. Conceptually, the basic goal-based model appears as follows:

Goal Achievement → Utility

On the other hand, as a general rule, it would also make sense to predict that:

Goal Achievement → Higher Individual Income

Thus, a correlation between income and utility would be anticipated. That relationship would only be causal, however, where competitive goals vis-à-vis peers versus self represented an important part of the individual's internal system of goals. Furthermore, the intensity of individual versus collectivist goals appears to vary considerably across cultures (Hofstede, 2001). Thus, we might also expect considerable variation in the impact of relative income on utility across cultures.

A second point may also be made regarding relative income utility models in economics. Achieving comparatively superior performance "with respect to a set of peer individuals and with respect to one's own past performance" is—in fact—a goal. In the illustration of Figure 4, for example, it seems closely related to the generic goal of "Winning". Thus, perhaps inadvertently, economists have adopted a goal-based utility model in this context, even if not described as such.

Perhaps the greatest difference between the neoclassical and goal-based utility models can be characterized in terms of opacity of decision processes. The neoclassical model is constructed very much according to the behaviorist philosophy in which a stimulus is provided in the form of a set of alternatives, response is determined in the form of a choice, and relatively little attention is paid to the decision maker's self-reported internal thought processes. The advantage of such an approach, particularly for macro-economic purposes, is that errors in predicting individual decisions are likely to get aggregated away. Moreover, as previously noted, models that assume utility results from satisfying drives—particularly Becker's (1976) economic model and evolutionary models (e.g., Gandolfi et al., 2002)—would tend to be relatively stable and therefore suitable for mathematical analysis and modeling. Weaknesses in neoclassical models only become embarrassingly apparent when specific individual preferences and behaviors become the object of prediction.

The specific goal-based model—presumed most active for decisions that are neither entirely ill-structured (employing generic goals and detectors) nor completely programmed (involving auto-

mated, unconscious task performance and goal progress monitoring)—presumes continual adoption, modification, completion, and abandonment of specific goals. It therefore complements the neoclassical model by being entirely decision-maker directed. It involves analyzing information processing activities within the decision maker's head, making it a cognitive rather than behavioral model. Indeed, without knowing the decision maker's individual goals, it would be nearly impotent in its ability to make any useful predictions. Its advantage would, in principal, therefore lie in interpreting behaviors and making predictions at the individual level. Since informing systems are often designed around a particular individual or known set of individuals, it would be particularly appropriate for analyzing utility in these contexts.

Implications for the Informing Sciences

In the context of the informing sciences, there are a number of potential implications of the proposed utility model. Foremost among these is its influence on the value of information. In economics, the value of a particular item is frequently viewed in terms of its marginal utility, which is to say that the value of an additional item of X_i to a particular individual, mathematically described as:

dU/dX_i

Ideally, an individual would want to spend each new dollar that becomes available on the good that offers the greatest marginal utility for that dollar. It therefore follows that a rational individual should choose a basket of goods and services such that each marginal utility is proportional to its price.

Since one example of a good or service that a client might choose to purchase is information relating to a particular task, we would expect that the price he or she would be willing to pay would be proportional to the perceived marginal utility of that information. So far, we have not left the domain of neoclassical economics.

The particular impact of the proposed model occurs when the nature of the informing taking place is considered. As illustrated in Figure 5, the types of tasks that may be associated with an informing system can vary significantly. On the one hand, a system could provide routine informing on a familiar task. For example, a system might provide up-to-the-minute vendor price quotes, used to inform purchasing decisions. For such a system, the neoclassical utility model might work perfectly well—by determining how valuable it is to have such quotes (as opposed to acquiring them through other means), you should be able to gauge how much the client would be willing to pay for the informing. Issues such as goals or drives would not play a particularly important role, since the utility function involved would already be well-learned and largely tied to task progress. Thus, the motivation for informing of this sort is likely to be almost task-driven.

At the other end of the continuum, the utility function would be entirely different and very much out of line with neoclassical models. Here, utility would derive from the client's perception that
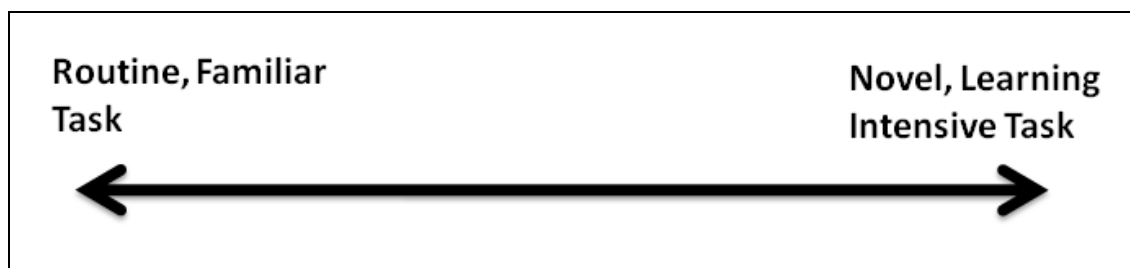

Routine, Familiar Task

Novel, Learning Intensive Task

Figure 5: Continuum of informing tasks

the informing will satisfy one of more of the generic drives and/or desires hard-wired into all of us. Such utility is likely to be highly vulnerable to framing effects and other cognitive biases. It will also be substantially impacted by visceral factors, such as emotions, which alter activation levels and attention associated with the various nodes. We would also expect the strength of these drives to vary considerably between individuals (Reiss, 2000). As a consequence, we would expect utility at this end of the continuum to be highly volatile and therefore largely client driven.

In between the two informing extremes, we have the region where motivation—and hence utility—is principally a product of making progress towards task-related goals. In this region, it follows that the source of the specific goals that are established—which could be the sender, the client (or the client's supervisor), or even the task itself—will have the greatest ability to impact the perceived value of informing. It would also seem to be the area where the greatest potential to influence the informing process exists. For example, the leader who is able to get the client to commit to certain goals that require informing will thereby increase the utility of the informing process to that client. Such leadership is likely to be far less effective at the Figure 5 extremes. Where the task is routine, such goals are no longer needed. Where the task is so unfamiliar to the client that utility is dominated by generic drives, variability across clients will make it much harder to impact utility in a predictable way.

Conclusions

In economics, utility is the function that drives value. Unfortunately, economic models of rational, utility-driven behavior tend to fail in unfamiliar settings and tend to ignore learning altogether. Such models, therefore, tend to be rather limited in their ability to say useful things about real-world informing. Only the most mechanical of informing systems functions with neither learning nor the unexpected.

Rather than abandon utility altogether, the present paper has attempted to construct a model of utility that synthesizes findings on learning and motivation from a wide range of disciplines, including management, economics, cognitive science, neuroscience, and connectionism. The model, intended to be cognitively plausible rather than mathematically convenient, proposes that utility passes through a continuum of stages, starting with utility that arises from a set of relatively generic drives that are hardwired into all of us and ending with utility that is, essentially, hardwired to progress in completing whatever task that the individual is focused on. The rough stages of this progress, as might be encountered for task performance in an informing system, are summarized in Table 3. Why we would need to understand these stages prior to applying utility to informing situations should be readily apparent; a utility model that fails to accommodate task unfamiliarity and learning processes would be sterile indeed for purposes of understanding the motivation to inform and be informed.

In addition to the obvious direct relevance of utility to informing, the paper also can be viewed as a case study on the challenges presented by excessive specialization and, particularly, the creation of disciplinary silos that do not talk to each other (Cohen, 1999). A principal source of the weaknesses in existing conceptions of utility is the concept's separate trajectory in many disciplines. In economics, it is called utility and has evolved so as to ensure the desired level of mathematical tractability and suitability for policy analysis; in the decision sciences, a virtual cottage industry has developed around demonstrating utility's axiomatic inconsistencies but such analysis is mainly limited to lab settings involving tasks unfamiliar to subjects; in applied psychology utility corresponds to motivation and it has been particularly influential in explaining non-economic factors influencing choice but, once again, mainly done so in controlled lab settings; in education, it is also aligned with motivation, but its focus is specialized on learning situations and entirely ignores economic aspects; in evolutionary biology, it is called fitness and it is explained largely in terms of characteristics promoting survival. Because these disciplines do not regularly talk to

each other, each has developed their own version of the concept that has strengths in the context of use, but also exhibits many weaknesses when considered outside of that context. An important role for a transdiscipline, such as the informing sciences, is to attempt to reconcile these separately evolved views into a more coherent whole.

Table 3: Utility sources for task performance

Source of Utility	Description
Task progress	Intrinsic utility derives from the process of completing the task. Because the task has become routine, the conscious satisfaction of individual task-related goals is no longer required. In this stage, neoclassical models of the utility of information are likely to be valid.
Task goals	Intrinsic utility derives from satisfying and progressing towards a series of specific goals that are presented as part of the task. These goals may serve to direct activities or act as targeted levels of achievement. In this stage, task-related utility is best predicted by the goal setting models so extensively documented in management and psychology. Inconsistencies in decision-making and biases will tend to lose impact in this region as goals become increasingly well established. As the task is repeated, specific goals and tradeoffs between goals become increasingly automated, ultimately leading to migration towards the task progress stage.
Generic drives and desires	Intrinsic utility derives from satisfying and progressing towards generic drives and desires that are present, to varying degrees, in all individuals. At this level, which is applicable mainly to highly unfamiliar and learning-oriented tasks, utility tends to be highly subject to framing issues and cognitive biases of the sort that economists and decision scientists have identified in experiments—often inconsistent with "rational" models of behavior. As task-specific goals are learned, the utility migrates towards the task goals stage.

References

Ainslie, G. (2001). *Breakdown of will.* Cambridge, UK: Cambridge University Press.

Allen, R.G.D. (1959). *Mathematical economics* (2nd ed.). London, UK: MacMillan.

Ariely, D. (2008). *Predictably irrational*. New York: Harper Collins.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*, 248-287.

Batson, C., & Shaw, L. (1991). Evidence for altruism: Towards a pluralism of prosocial motives. *Psychological Inquiry, 2*(2), 107-122.

Becker, G. (1976). *The economic approach to human behavior.* Chicago: University of Chicago Press.

Bell, D. (1979). Multiattribute utility functions: Decompositions using interpolation. *Management Science, 25*(8), 744-753.

Bell, D., Raiffa, H., & Tversky, A. (1988). Descriptive, normative, and prescriptive interactions in decision making. In D. Bell, H. Raiffa, & A. Tversky, *Decision making: Descriptive, normative, and prescriptive interactions* (pp. 9-30). Cambridge, UK: Cambridge University Press.

Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness and utility: An explanation for the Easterlin Paradox and other puzzles. *Journal of Economic Literature, 46*(1), 95-144.

Cohen, E. (1999). Reconceptualizing information systems as a field of the transdiscipline informing science: From ugly duckling to swan. *Journal of Computing and Information Technology, 7*(3), 213-219.

Dijksterhuis, A., Bos, M., Nordgren, L., & van Baaren, R. (2006). On making the right choice: The deliberation-without-attention effect. *Science, 311,* 1005-1007.

Dyer, J. (1972). Interactive goal programming. *Management Science, 19*(1), 62-70.

Elster, J., & Loewenstein, G. (1992). Utility from memory and anticipation. In G. Loewenstein & J. Elster, *Choice over time* (pp. 213-234). New York: Russell Sage Foundation.

Forster, J., Higgins, T., & Idson, L. (1998). Approach and avoidance strength during goal attainment: Regulatory focus and the 'goals loom larger' effect. *Journal of Personality and Social Psychology, 75*(5), 1115-1131.

Frey, D., Kumpf, M., Irle, M., & Gniech, G. (1984). Re-evaluation of decision alternatives dependent upon the reversibility of a decision and the passage of time. *European Journal of Social Psychology, 14*, 447-450.

Gandolfi, A. E., Gandolfi, A. S., & Barash, D. (2002). *Economics as an evolutionary science: From utility to fitness*. New Brunswick, NJ: Transaction Publishers.

Gilbert, D. (2007). *Stumbling on happiness*. New York: Knopf.

Gill, T. G. (1996). Expert systems: Task change and intrinsic motivation. *MIS Quarterly, 20*(3), 301-329.

Gill, T. G. & Hicks, R. (2006). Task complexity and informing science: A synthesis. *Informing Science: the International Journal of an Emerging Transdiscipline, 9*, 1-30. Retrieved from http://inform.nu/Articles/Vol9/v9p001-030Gill46.pdf

Gladwell, M. (2005). *Blink*. New York: Back Bay Books.

Grant, H., & Dweck, C. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology, 85*(3), 541–553.

Grether, D., & Plott, C. (1979). Economic theory of choice and the preference reversal phenomenon. In S. Lichtenstein & P. Slovic, (2006). *The construction of preference* (pp. 77-93). Cambridge, UK: Cambridge University Press.

Hackman, J. R., & Oldham, G. R. (1980). *Work redesign*. Reading, MA: Addison Wesley.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Thousand Oaks, CA: Sage Publications.

Holland, J. H. (1995). *Hidden order*. Reading, MA: Helix Books.

Janssen, O., & Van Yperen, N. (2004). Employees' goal orientations, the quality of leader-member exchange, and the outcomes of job performance and job satisfaction. *Academy of Management Journal, 47*(3), 368–384.

Johnson, R., Chang, C., & Lord, R. (2006). Moving from cognition to behavior: What the research says. *Psychological Bulletin, 132*(3), 381–415.

Kahneman, D., Slovic, P., & Tversky, A. (Ed.) (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-292.

Keeney, R. (1972). Utility functions for multiattributed consequences. *Management Science, 18*(5), 276-287.

Keeney, R. (1988). Value-focused thinking and the study of values. In D. Bell, H. Raiffa, & A. Tversky, *Decision making: Descriptive, normative, and prescriptive interactions* (pp. 465-494)*. Cambridge, UK: Cambridge University Press.

Klein, H. (1989). An integrated control theory model of work motivation. *Academy of Management Review, 14*(2), 150-172.

Klein, H., Wesson, M., Hollenbeck, J., & Alge, B. (1999). Goal commitment and the goal setting process: Conceptual clarification and empirical synthesis. *Journal of Applied Psychology, 84*(4), 885-895.

Langer, E. (1975). The illusion of control. In D. Kahneman, P. Slovic, & A. Tversky, (1982), *Judgment under uncertainty: Heuristics and biases* (pp. 231-238)*. Cambridge, UK: Cambridge University Press.

Latham, G., & Locke, E. (2006). Enhancing the benefits and overcoming the pitfalls of goal setting. *Organizational Dynamics, 35*(4), 332–340.

Latham, G., & Steele, T. (1983). The motivational effects of participation versus goal setting on performance. *Academy of Management Journal, 26,* 406-417.

Lawrence, P. R., & Nohria, N. (2002). *Driven: How human nature shapes our choices.* San Francisco, CA: Josie-Bass.

Lee, F., Sheldon, K., & Turban, D. (2003). Personality and the goal-striving process: The influence of achievement goal patterns, goal level, and mental focus on performance and enjoyment. *Journal of Applied Psychology, 88*(2), 256–265.

Locke, E. (1978). The ubiquity of the technique of goal setting in theories of and approaches to employee motivation. *Academy of Management Review, 3*, 594-601.

Locke, E. (2004). Goal-setting theory and its applications to the world of business. *Academy of Management Executive, 18*(4), 124-125.

Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. In G. Loewenstein & J. Elster, *Choice over time* (pp. 119-146). New York: Russell Sage Foundation.

Lopes, L., & Oden, G. (1999). The role of aspiration level in risky choice: A comparison of cumulative prospect theory. *Journal of Mathematical Psychology, 43*, 286-313.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review, 88*(5), 375-407.

Moschandreas, M. (1997). The role of opportunism in transaction cost economics. J*ournal of Economic Issues, 31*(1), 39-57.

Newell, A. (1990). *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

Nunes, J., & Drèze, X. (2006). The endowed progress effect: How artificial advancement increases effort. *Journal of Consumer Research, 32*(4), 504-512.

Payne, J., Bettman, J., & Johnson, E. (1993). *The adaptive decision maker.* Cambridge, UK: Cambridge University Press.

Politser, P. (2008). *Neuroeconomcs.* Oxford, UK: Oxford University Press.

Read, D. (2004) Intertemporal choice. In D. Koehler & N. Harvey, *Blackwell handbook of judgment and decision making* (pp. 424-443). Malden, MA: Blackwell.

Reiss, S. (2000). *Who am I? The sixteen basic desires that motivate our actions and define our personalities.* New York, NY: Berkley Books.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review, 89*(1), 60-94.

Selfridge, O.G. (1958). Pandemonium: A paradigm for learning. In J. A. Anderson & E. Rosefield, (1988). *Neurocomputing: Foundations of research* (pp. 117-122). Cambridge, MA: MIT Press.

Shermer, M. (2008). *The mind of the market: Compassionate apes, competitive humans and other tales from evolutionary economics*. New York, NY: Times Books.

Shiffrin, R. M., & Dumais, S. T. (1981). The development of automatism. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 111-140). Hillsdale, NJ: Lawrence Earlbaum.

Simon, H. (1976). *Administrative behavior* (3rd ed.). New York, NY: Free Press.

Simon, H. (1981). *The sciences of the artificial* (2nd ed.). Cambridge, MA: MIT Press.

Simon, H. (1991). Organizations and markets. *Journal of Economic Perspectives, 5*(2), 25-44.

Simon, H. (1992). Altruism and economics. *Eastern Economic Journal, 18*(1), 73-83.

Slovic, P., Finucane, M. L., Peters, E., & Macgregor, D. G. (2006). The affect heuristic. In S. Lichtenstein & P. Slovic, *The construction of preference* (pp. 434-453). Cambridge, UK: Cambridge University Press.

Snyder, C. (2002). Hope theory: Rainbows in the mind. *Psychological Inquiry, 13*(4), 249-275.

Soman, D. (2004). Framing, loss aversion, and mental accounting. In D. J. Koehler & N. Harvey,(Eds.), *Blackwell handbook of judgment and decision making* (pp. 379-398). Malden, MA: Blackwell.

Stajkovic, A., Locke, E., & Blair, E. (2006). A first examination of the relationships between primed subconscious goals, assigned conscious goals, and task performance. *Journal of Applied Psychology*, *91*(5), 1172–1180.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. In D. Kahneman, P. Slovic, & A. Tversky, *Judgment under uncertainty: Heuristics and biases* (pp. 163-178). Cambridge, UK: Cambridge University Press.

Tversky, A., & Kahneman, D. (1988). Rational choice and the framing of decisions. In D. E. Bell, A. Raiffa, & A. Tversky (Eds.), *Decision making: Descriptive, normative and prescriptive interactions* (pp. 167-192). Cambridge, UK: Cambridge University Press.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108*(3), 550-592.

Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychological Review, 111*(3), 757-769.

Vroom, V H. (1964). *Work and motivation.* New York: Wiley.

Wilson, T., & Schooler, J. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology, 60*(2), 181-192.

Wofford, J., Goodwin, V., & Premack, S. (1992). Meta-analysis of the antecedents of personal goal level and of the antecedents and consequences of goal commitment. *Journal of Management, 18*(3), 595-615.

Biography

Grandon Gill is an Associate Professor in the Information Systems and Decision Sciences department at the University of South Florida. He holds a doctorate in Management Information Systems from Harvard Business School, where he also received his M.B.A. His principal research areas are the impacts of complexity on decision-making and IS education, and he has published many articles describing how technologies and innovative pedagogies can be combined to increase the effectiveness of teaching across a broad range of IS topics. Currently, he is an Editor of the *Journal of IT Education* and an Associate Editor for the *Decision Science Journal of Innovative Education.*