Predicting and Indulging Changing Preferences

GEORGE LOEWENSTEIN and ERIK ANGNER

The one thing I could be sure of was that I had to leave this apartment, where I had never known a moment's peace of mind, as soon as possible. . . . The trouble, the rub, was that I had to give three months' notice and therefore had to predict how I would be feeling three months hence, which was very difficult. It was all very well deciding today that I wanted to leave but what counted was how I was going to be feeling three months from now. You could be perfectly happy today, I would say to myself, and three months from now you could be suicidal, precisely because you will see the enormity of the mistake you made months earlier.

—Geoff Dyer, Out of Sheer Rage

Decisions, from the most mundane to the most momentous, often involve a prediction of future preferences. Whether one is shopping for groceries, contemplating whether to "tie" the knot, or (as in the epigraph) deciding whether to sign the lease on an apartment, the feelings and tastes that matter may not be those one currently has but rather those that one anticipates having when the consequences of the decision are experienced. Mispredicting future preferences can result in diverse negative consequences, from uneaten groceries to painful divorces to the suicidal feeling of being trapped for another year in an apartment one detests.

The trouble—as Dyer puts it, the rub—is that predicting preferences is difficult. The difficulties stem in part from the fact that determinants of tastes are complex and poorly understood. Though social scientists have devoted whole careers to studying the formation of
preferences, progress has been painfully slow. To date there is little agreement about the sources of even the most fundamental preferences—for example, for food, drink, and sex. Given the lack of progress by social scientists, the fact—documented by research reviewed here—that people have trouble predicting how their preferences will change should come as no surprise.

The difficulties people have in predicting their future preferences are exacerbated by the fact that to imagine having tastes that are substantially different from those we have at the moment is challenging. We have a tendency to believe that our current tastes and preferences reflect objective features of the external world to a larger extent than they actually do. As Adam Smith (2000 [1759], 283) observed almost two hundred and fifty years ago,

> Few men... are willing to allow, that custom or fashion have much influence upon their judgments concerning what is beautiful, or otherwise. [Rather, they] imagine that all the rules which they think ought to be observed... are founded upon reason and nature, not upon habit or prejudice.

Ross and Ward (1996) use the term *naïve realism* to denote this tendency to believe that our perceptions and tastes are more objective, and hence more universal, than they really are.' As a result of our susceptibility to naïve realism, current preferences for cars, clothes, music—even body types and intimate relations—simply seem “right,” and it is difficult to imagine that were we in a different culture or historical period our preferences might be quite different.

Predicting changes in future preferences is also difficult because tastes and desires are an integral aspect of our personal *identity* (Akerlof and Kranton 2000; Belk 1988; Frederick, chap. 2 herein; Parfit 1971, 1982). People define who they are in part by their tastes and values; imagining oneself with different preferences is therefore similar to imagining oneself as a different person, which is a difficult mental exercise.

Owing to difficulties involved in predicting preferences, mistakes are common (or so we later argue, reviewing available research on the topic). The difficulties just noted here suggest that one kind of mistake will be particularly common: predictions will be too “regressive,”—that is, biased in the direction of current tastes. For example, if people are unaware of certain sources of preference change, they are likely to underestimate the magnitude of changes caused by that source. Similarly, if people have difficulty imagining having preferences different from their current ones—either because they mistakenly view their current tastes as objective or perceive their current tastes as an integral part of their core identity—then they will tend to underestimate the magnitude of those changes. Indeed research suggests that people are prone to exactly such a regressive bias, referred to by Loewenstein, O’Donoghue, and Rabin (2001) as *projection bias*.

Analyses of shifting preferences by economists and decision researchers typically assume that people want to satisfy whatever preferences they expect to hold at the time when the consequences of their decision are enjoyed or suffered (whether or not such analyses acknowledge that people sometimes mispredict their future preferences). This assumption is implicit, for example, in economic models of “habit formation” (Duesenberry 1952; Pollak 1970), including models of addiction (for example, Becker and Murphy 1988). However, the assumption that people attempt to honor future preferences is not always valid. In some cases, people indeed attempt to deny their anticipated future preferences. For example, in “cool” moments, people recognize that they may get “hot” in the future and develop transient preferences for things they currently would prefer to avoid (for example, harmful drugs, dangerous sexual practices, or unhealthy foods). To prevent themselves from behaving in this fashion, people sometimes seek to deter themselves from acting on their anticipated future preferences (for example, by taking antabuse, which makes alcohol intake nauseating) or simply remove the undesired behavior from their future choice set (for example, keeping alcohol or tempting snack food out of the house; see Elster 1979; Schelling 1984).

Instead of simply trying to honor the preferences they have or expect to have, people also sometimes attempt to shape their own tastes. Rather than attempting to prevent themselves from succumbing to hot future preferences, for example, people sometimes avoid exposing themselves to situations and stimuli that could cause them to become hot, thereby saving themselves from experiencing a transient preference that they do not wish to have. People also sometimes attempt to refine their future tastes—for example, they drink fine wines or listen to highbrow music that they don’t currently enjoy in the hope of developing a taste for it. To date, very little research has sought to understand the factors that cause people to indulge, deny, or seek to change their own future preferences. Instead most of the attention to this issue has come from philosophers. A later section summarizes this work and attempts to draw out some general conclusions about when people in fact choose to indulge or deny anticipated changes in their tastes. In passing, we also discuss some related normative is
sues, namely, under what conditions people should indulge or deny anticipated preferences.

Beyond discussing the twin problems of predicting and honoring tastes, this chapter has a third goal, which is to enumerate some of the diverse determinants of preferences and the sources of changes in tastes. This goal is in certain ways more basic, since whether people manage to accurately predict, and whether they choose to honor (or should choose to honor), future preferences may depend on the source of those preferences. Thus an understanding of the determinants of preferences is critical to understanding when people mispredict changes in their own preferences and when they choose to honor or deny such changed preferences.

**Sources of Preference Change**

**Do Preferences Change at All?**

In their famous article “De Gustibus Non Est Disputandum,” Stigler and Becker (1977, 76) argue that tastes and preferences are stable over time and are identical across people; they write, “tastes neither change capriciously nor differ importantly between people . . . one does not argue over tastes for the same reason that one does not argue over the Rocky Mountains—both are there, will be there next year, too, and are the same to all men.” This seemingly obvious false assertion depends crucially on the distinction between two kinds of entities: final preferences (or demand) and underlying preferences (or tastes). Although final preferences may vary, say, when a new fashion appears or when one gets sick of spam after eating too much of it, one can usually identify some set of stable underlying preferences—for example, looking good and experiencing the pleasures of the palate.

Herbert Simon (1981, 58) explicitly challenged the perspective advanced by Stigler and Becker, arguing that the distinction between final and underlying preferences obscures more than it illuminates.

It is unrealistic to suppose that utility functions are given and remain fixed. New experiences produce new tastes. Some attempts have been made to save the classical theory along this dimension by replacing tangible goods and services as the arguments of the utility function with more basic “wants”—for example, pleasure from music listened to, rather than number of hours of listening. Thus, Becker and Stigler speak of investing in musical experience to increase the pleasure, per unit of time, in listening to music.

It may be doubted whether anything is gained by trying to rescue the traditional view of utility with such heroic measures. If, to continue the example, we do not wish to speak of a change in utility function as the result of listening to music, then we must postulate within the human head a production function (itself changeable by experience) that manufactures musical pleasure from musical listening. We have merely relocated “taste” from the utility function to that hypothetical new production function. It would seem more parsimonious simply to regard the utility function as an evolving structure.

Clearly, the argument between Stigler and Becker and Simon raises important questions about the proper definition of tastes and taste change. Should momentary changes in behavior—such as those that result from elevated appetites or emotions—be considered to reflect a change in tastes? What about behavior that reflects new information about engaging in some particular activity? Does a friend’s suggestion to avoid a certain movie produce a change in tastes, or simply provide us with information that allows us to predict our tastes more accurately? As Stigler and Becker imply, a meaningful distinction exists between fundamental changes in tastes and changes in demand. Our own intellectual proclivities, however, lie more with Simon than with Stigler and Becker. Defining away individual differences and intraindividual changes in preferences only obfuscates the study of taste formation. The problem of predicting and honoring changing preferences remains whether or not we assume that every change in preferences rests on a foundation of stable underlying tastes. On Stigler and Becker’s account, predicting such preferences (assuming fundamental tastes are known) boils down to predicting the characteristics of the evolving “production function” mentioned by Simon. We adhere to the more customary terminology, while acknowledging the nontrivial question raised by Stigler and Becker of what forms of changes in demand should rightly be classified as changes in “preferences.”

Here we enumerate a number of different determinants of preference, and therefore sources of preference change (see McCauley, Rozin, and Schwartz 2002 for a much more complete treatment). We draw connections between those determinants, by grouping them into broad categories and showing how some of these categories can be modeled using similar mathematical formulations.
Endogenous Change in Tastes: Habit Formation, Satiation, and Refinement

"Endogenous change in tastes" (Hammond 1976) refers to a situation in which what one consumes in the present alters the preferences one has in the future. Within this broad category are several significantly different variants that can be distinguished largely in terms of whether past consumption increases or decreases one's preference for future consumption.

Habit Formation Habit formation refers to a situation in which consuming a particular substance increases one's preference for it. If \( s_t \) is the "habit stock" that summarizes the extent of past consumption of \( c \) (with higher values of \( s \) corresponding to greater past consumption) and the individual's utility function takes the form \( u(c, s) \), then habit formation corresponds to the case in which the marginal utility of \( c \) is increasing in \( s \)—that is,

\[
\frac{\partial^2 u(c, s)}{\partial c \partial s} > 0.
\]

A common assumption in modeling habit formation (for example, Ryder and Heal 1973, or see Frederick et al., chap. 1 herein) is that the habit stock shifts according to \( s_t = \alpha s_{t-1} + (1 - \alpha) c_t \). At one extreme, when \( \alpha \) equals 1, then there is no adaptation; the habit stock always remains at whatever level it began. At the opposite extreme, when \( \alpha \) equals zero, then the habit stock adjusts instantly to the level of consumption in the current period. For intermediate values of \( \alpha \), this formula implies that \( s_t \) is an exponentially weighted sum of past consumption, with more recent consumption given greater weight.

Within the category of habit formation, a distinction is often made between negative and positive habits. Negative habits—such as harmful addictions—arise when the "habit stock" \( s \) has a negative impact on overall utility—that is,

\[
\frac{\partial u(c, s)}{\partial s} < 0.
\]

Positive habit formation refers to situations in which enhanced liking of certain goods and activities such as good music, fine wine, and so on is thought to enhance overall utility—that is,

\[
\frac{\partial u(c, s)}{\partial s} > 0.
\]

Satiation Satiation can be thought of as the opposite of habit formation: it is a situation in which consuming more of a substance decreases its marginal utility. Satiation can be a short-term phenomenon. For example, after consuming a steak for dinner, one's desire for a second steak might be minimal or even negative. Yet satiation can also be a longer-term phenomenon, as might occur after consuming steak several evenings in a row, or after spending too many summer vacations in the same spot. Satiation is commonly represented with a simple utility function that incorporates diminishing marginal utility—that is, \( u(c) \), with \( u''(c) < 0 \). Yet such a formulation is only capable of dealing with the most short-term variant of the phenomenon—that is, the effect of steak just consumed on the marginal utility of further consumption of steak; it cannot deal with, for example, the effect of steak eaten on the previous day on the marginal utility of steak today.

More complex patterns of satiation can be modeled parsimoniously using the same state-dependent utility function that is commonly applied to habit formation. Satiation simply corresponds to the case in which marginal utility is a declining function of the habit stock:

\[
\frac{\partial^2 u(c, s)}{\partial c \partial s} < 0,
\]

the opposite of the pattern that characterizes habit formation. With such a formulation, to model the complexities of satiation—for example, by again assuming that \( s_t \) is a weighted average of past consumption—is easy. Since satiation diminishes some sources of pleasure without increasing others, it is probably safe to assume that satiation decreases (or at least does not increase) an individual's overall utility—that is, that

\[
\frac{\partial u(c, s)}{\partial s} \leq 0.
\]

Refinement In addition to habit formation and satiation, it is useful to distinguish a third pattern of endogenous preference change that could be termed refinement. People often expose themselves to goods and experiences for the purpose of refining their tastes. There are various possible interpretations of what refinement involves, but perhaps the most common is that it involves an increase in one's relative appreciation for higher-quality goods or experiences. If \( r \) represents the quality level of a particular type of good, then refinement would resemble habit formation in the sense that (though note that this ex-
pression substitutes $x$—the level of quality—for $s$, the habit stock. The big difference between positive and negative habit formation on the one hand, and refinement on the other, is in the impact of the habit on overall utility. Presumably, refining one's tastes not only increases one's enjoyment of high-quality goods but also, as illustrated in figure 12.1, decreases one's enjoyment of low-quality goods.

$$\frac{\partial^2 u(c,x)}{\partial c \partial x} = 0$$

If this is the case, then whether refining one's tastes actually increases overall pleasure will depend on any direct effect of refinement on utility, and on one's budget constraint—that is, on whether one can actually afford the goods for which one's appreciation has been enhanced. If the higher-quality goods one has developed a taste for are unattainably expensive, forcing one to consume goods whose inferiority is now recognized, then refining preferences may not increase utility. Having refined tastes in and of itself may be a source of utility, but certainly it cannot be assumed that having such tastes increases one's overall pleasure from consumption.\(^3\)

**Appetites, Emotions, and Other "Visceral Factors"**

Tastes are also determined at least in part by biological systems that fluctuate, sometimes dramatically, even over short periods of time. Like other animals, humans are sustained by the operation of numerous homeostatic processes. These processes regulate body temperature, blood pressure, food intake, heart rate, and a whole range of chemical and electrical processes in the body and brain. Homeostatic processes operate by comparing the level of a system that is being regulated (for example, blood oxygenation) to a set-point—a desired level. When the actual level of the system departs from the set-point, this triggers processes that shift the system in the direction of the set-point.\(^4\)

A wide range of regulatory processes are devoted to maintaining homeostasis in these diverse systems. Some of these processes are automatic and occur without conscious intention or even knowledge. For example, body temperature is maintained in part by autonomous mechanisms such as sweating; but other processes involve overt, deliberate behaviors. Thus body temperature is also maintained by actions such as turning on the air conditioner or heater or drinking a cold or hot beverage.

People typically are motivated to take these types of homeostasis-

![Figure 12.1 Refining Tastes](image)

*Source: Authors’ configuration.*

restoring actions through the combined application of a carrot and stick. The stick takes the form of specific discomforts, such as hunger, thirst, and pain that motivate us to take action. The carrot is the increase in pleasure that accompanies such actions. Both effects motivate us to take actions to bring the system back into line. Thus a decrease in our core body temperature produces intense discomfort, and any actions that can increase one's body temperature, such as placing one's hand in a bowl of warm water or drinking a hot beverage, become pleasurable. Our brains in effect produce transient changes in preferences that induce us to take actions to reinstate homeostasis. These combined effects can be modeled with the same state-dependent preferences used to model endogenous change in tastes. If $s_1$ represents the level of a particular homeostatic mechanism such as hunger, then the carrot corresponds to

$$\frac{\partial^2 u(c,s_1)}{\partial c \partial s_1} > 0,$$

and the stick to

$$\frac{\partial u(c,s_1)}{\partial s_1} \leq 0.$$

Note that these are the same partial derivatives that define a negative habit. Moreover, as some goods and activities become more attractive, others become less attractive, presumably so that the organ-
ism does not waste effort pursuing goals with a lower priority for survival (Brendl, Markman, and Irwin 2001).

The cues that signal departures from homeostatic set-points in some cases can be extraordinarily sophisticated. Thus, for example, signals to the brain that trigger hunger come from the stomach, intestines, blood-sugar level, and so on. Moreover, we are sensitive not only to internal signals but also to external signals that indicate the possible future scarcity of food, which may help to explain why dieting is so ineffective (see Herman and Polivy, chap. 16 herein). Perhaps when a person begins to diet, the brain receives signals indicating that the body is about to be deprived of food and responds by increasing the person’s motivation to eat. The brain is also attuned to opportunities: if it senses an opportunity for low-cost access to food, it again responds by increasing the motivation to eat (again, causing problems for the dieter).

Conditioning

Classical conditioning is one of the most coherent and well-researched sources of preference change. The idea of classical conditioning begins with the basic building blocks of “unconditioned stimuli” (US), which are rewards and punishments, such as food, water, and electrical shock, that we are biologically programmed to find attractive or repellant. During development, organisms, including humans, come to associate certain types of cues with unconditioned stimuli, and these cues become “conditioned stimuli” (CS) that produce reactions similar to those of the US with which they are associated. For example, a baby might enjoy the taste of mother’s milk, and when the mother rapidly becomes a cue signaling the availability of milk, it transfers the same positive feelings toward the mother (of course maternal attention may itself also be an unconditioned stimulus). The process need not end here. The individual who has become conditioned to like milk might subsequently develop positive feelings toward locations (for example, rooms) where he or she often consumes milk—a process known as conditioned place preference. Thus classical or “evaluative” conditioning (another common label for the process of acquiring preferences via associative connections) can explain how initial fundamental likes and dislikes—unconditioned stimuli—can transfer their valence to a wide range of objects, persons, and activities in an ever-widening web.

Evaluative conditioning has been observed in diverse domains and with diverse stimuli (see De Houwer, Thomas, and Baeyens 2001 for a recent review). For example, evaluations of political slogans and human faces have been enhanced or undermined by pairing these CS with pleasant or aversive odors; drink flavors have been enhanced by their pairing with sweetness (that persists when the sweetness is later eliminated); and neutral pictures have been made to appear attractive or unattractive by presenting them simultaneously with pictures that were selected to be attractive or unattractive.

Different conditioned reactions seem to differ in terms of how quickly they are acquired, how rapidly they are lost when not reinforced, and a variety of other properties. Thus fear conditioning and learned taste aversion (in which a person or other type of animal avoids food that has become associated with nausea) occur very rapidly—often in a single trial—whereas positive conditioning effects usually require repeated pairings. Some forms of conditioning seem to be very specific to the environmental settings in which they are learned, while others are not. Different forms of conditioning differ in how rapidly they become “extinguished” when the conditioned stimulus is presented in the absence of the unconditioned stimulus (for further nuances concerning extinction see Bouton 1994; Bouton and Swartzentruber 1991; LeDoux 1996). Moreover, different species, including humans, seem to be biologically “prepared” to learn some types of CS, US associations (for example, between a snake or angry face and a shock) but not others (for example, between a flower or rabbit and a shock) (see, for example, Öhman 1986).

Conditioning effects do not require conscious awareness of US–CS associations. In one study, subjects were shown a series of pictures of a target person engaged in various activities, such as getting into a car and grocery shopping. There were two experimental conditions. In both, prior to the presentation of each photograph, another photograph was flashed subliminally (so rapidly that it was not consciously perceived). In one condition, those photographs were intended to arouse positive affect (for example, kittens or smiling friends), and in the other they were intended to arouse negative affect (for example, a skull, face on fire, bucket of snakes). At the end of the session, subjects were asked a series of questions about how much they liked or disliked the person in the photo. Those who had been exposed to the negative subliminal photos expressed substantially less liking than those exposed to the positive subliminal photos. A second component of the same study showed further that subjects made sense of their own tastes by evaluating various personality attributes of the target person differently. That is, people seemed to seek out rational explanations for why they either liked or disliked the individual shown in the photos, and were of course completely unaware that their like or dislike was in fact powerfully shaped by subliminal conditioning. In
general, because people tend to rationalize their tastes and because conditioning is usually an unconscious process, the natural tendency is to underestimate the impact of conditioning—to believe that tastes are, as Adam Smith expressed it, "founded upon reason and nature."

In sum, numerous studies have established the occurrence and robustness of evaluative conditioning effects, and many researchers have asserted the importance of conditioning for processes such as the formation of food tastes, the acquisition of phobias and of sexual fetishes (for example, Bouton, Mineka, and Barlow forthcoming). Yet few if any studies have produced a "smoking gun"—that is, have convincingly demonstrated the ecological importance of conditioning as a determinant of tastes. Indeed there have been some failures to observe evaluative conditioning in situations in which one might have strong expectations that it would occur. Heroin addicts, for example, in treatment whose methadone is diluted in a solution of TANG do not subsequently develop a taste for TANG, even though relief of heroin craving would seem to be an extremely positive stimulus (see Rozin and Zellner 1985, 195).

**Maturation**

Maturation can be defined as a situation in which preferences change in a systematic fashion as a function of time:

\[
\frac{\partial u(t)}{\partial t} \neq 0.
\]

Although these changes are most dramatic in infancy and early childhood, when nutritional needs and preferences develop in a highly predictable manner, they continue into adulthood. For example, puberty brings about significant changes in sexual and other preferences. Similarly, young adults often acquire tastes for substances such as coffee, whiskey, and cigarettes, and shed other preferences, such as that for extremely sweet breakfast cereal. That younger people tend to be more radical and older people more conservative may also be true, though the evidence on this point is ambiguous (see Glenn 1980, 619ff).

The field of marketing, for obvious reasons, is vitally concerned with the question of how tastes and consumption patterns change with age, and marketing researchers have observed an especially interesting pattern of maturation for certain kinds of preference. Holbrook and Schindler (1989) found that adults tend to like whatever music was popular when they were in their mid-twenties, and that preferences in music tended to decline in a tentlike fashion around the music of this period. In more recent studies, the researchers found closely related patterns in preferences for fashion models (Schindler and Holbrook 1993), photographs of movie stars (Holbrook and Schindler 1994), and motion pictures (Holbrook and Schindler 1996). Despite the similarity in qualitative pattern, however, the specific age at which preferences peak appears to vary by stimulus. For popular music, that age is roughly twenty-four, for fashion models thirty-three, for pictures of movie stars fourteen, and for motion pictures twenty-seven (Holbrook and Schindler 1996, 34). Thus different types of preferences tend to become "frozen" at different critical periods in one's life.

A similar pattern has emerged in research on political preferences. In particular, party identification tends to be fairly stable during one's adult years (Green and Palmquist 1990, 1994). Many researchers see this as evidence that "once partisan or other social identifications take root in young adulthood, they tend to persist even amid changing political circumstances" (Gerber and Green 1998, 795). The traditional explanation suggests that party identification becomes part of one's self-concept, which serves as an impediment to subsequent change. Many of the preference changes that take place over a lifetime may be due to physiological effects associated with aging. For example, there is evidence that the sense of smell is transformed over time. Seifert and colleagues (1997, 595) report that the ability to recognize a wide variety of foods decreases over the adult years. Similarly, Doty and colleagues (1984, 1441) conclude that the "average ability to identify odors reaches a peak in the third and fourth decades of life, and begins to decline monotonically after this time." This, they argue, explains why elderly people often complain that their food lacks flavor (Doty et al. 1984, 1443), and it could explain why older people are more tolerant of foods with an unpleasant smell (Pelchat 2000).

Maturation and aging affect preferences in part indirectly, by changing our susceptibility to external influences. Again this is particularly clear in children, where younger children aged three to four are far more vulnerable to suggestion than are older children aged five to six (Ceci and Huffman 1997). Yet the progressive stabilization of preferences with aging seems to continue well into adulthood, as the research of Holbrook and Schindler confirms. Glenn (1980, 602) finds that research largely confirms the truism that "attitudes, values and beliefs tend to stabilize and become less likely to change as persons grow older." For better or worse, over time we appear to become increasingly inoculated against external influences.
Social Influence

Social factors clearly have a powerful influence on tastes. As attested to by international differences and historical changes, people are capable of forming diverse tastes for food, attire, architecture, styles of furniture, manners, and so on. While good functional reasons may exist for why certain tastes become widespread in certain cultures at certain points, no doubt social transmission is the main mechanism by which such tastes are disseminated.

Social influences can be as subtle as they are powerful. The literature on co-action effects indicates that eating and drinking behavior is strongly affected by the presence of others, though few people seem to be aware of the effect. Watson and Sobell (1982) found that males participating in a beer taste test drank significantly more—on average about twice as much—when paired with a heavily drinking companion than when paired with a companion who was not drinking at all. Roth (1999, ii) reports that subjects ate fewer cookies when in the presence of a non-eating observer than when alone, and that they tried to match the intake of the companion when both were eating. Roth explains observed behavior as the result of two social norms—one in favor of minimal eating and one in favor of matching the food intake of the other—ultimately driven by a concern for impression management (Roth 1999, 114ff.). Animals show co-action effects as well. Zajonc (1965) reviews research showing that rats, chickens, and puppies eat significantly more when coupled with other hungry individuals. An apparently fully sated chicken, he reports, will eat up to two-thirds as much again when introduced to a hungry companion chicken.

Though a tendency to conform can lead to erroneous judgments, overeating, and so on, conformity can also be good for you. Deutsch and Gerard (1955, 629) identify two important functions that may be served by conforming. First, conformity may satisfy a desire to live up to the expectations of others. Second, conformity may be a useful heuristic in cases of uncertainty about fact or value. Economists recently have become interested in social influence and have formulated models that incorporate both of these effects. For example, Gale (1996) proposes a model of herd behavior, in which people imitate others because they provide valuable information about the utility of actions. Models of conspicuous consumption (Veblen 1899) see consumption as driven in part by a desire to project a favorable image and to achieve social status (Frank 1985; Corneo and Jeanne 1997). Bernheim (1994) develops a model in this spirit in which agents care not only about “intrinsic” utility (that is, utility they derive directly from consumption) but also their social status, which they take to be influenced in part by their market choices. The model implies conformity, since agents “recognize that even small departures from the social norm will seriously impair their status” (Bernheim 1994, 841).

To say how exactly the social context affects behavior is often difficult. If you invite your friends for dinner but decline to drink wine yourself, they will likely drink less than they otherwise would have. Why is that? Perhaps they think you will not like it if they become uninhibited when you remain sober. Perhaps they assume some social norm that prohibits drinking in this situation. Perhaps—knowingly or not—they adopt what they take to be your current attitude in favor of sobriety. Perhaps they infer that something may be wrong with the wine. As the example suggests, social influences on consumption commonly operate through a mixture of mechanisms, some but perhaps not all of which should properly be counted as causing changes in tastes.

Motivated Taste Change

The famous song “Love the One You’re With” performed by the 1960s rock group Crosby Stills Nash and Young highlights the benefits of motivated taste change. If one could only learn to love what one has, or what no one else wants, most people would lead exceptionally fulfilled lives. Unfortunately, humans are not generally constituted in this fashion (which is why we need songs to encourage us to try nevertheless). Quite the opposite, as suggested by research on social influences, we often seem predisposed to want exactly what everyone else wants. This means that we often value what is scarce and expensive.

There are probably good evolutionary reasons for the limited sway we seem to have over our preferences. If people could make their preferences conform to their current attainments, then humans—for better or worse—would have undoubtedly made a lot less progress during our brief time on the planet.

Nevertheless, evidence shows that in some cases people can shift their preferences in favor of what they possess. The oldest demonstrations of what seems to be motivated taste change were produced by cognitive dissonance researchers. Cognitive dissonance was hypothesized to be a negative affective state experienced when one’s beliefs are inconsistent with one another or when actions are out of line with beliefs or preferences. Dissonance researchers believed that avoidance or elimination of dissonance motivated people to shift their beliefs or preferences into line with one another. Consistent with their predic-
ions, dissonance researchers demonstrated that people tended to increase their preference for objects they had chosen and to decrease their preference for objects they had rejected (for example, Brehm 1956; Festinger 1964).

Dissonance researchers postulated and found that changes in preference occurred only when people had freely chosen objects, and not when the objects were simply given to them. Subsequent research, however, has found similar effects on liking even when ownership did not depend on subjects’ decisions. Beggan (1992), for example, found that simply possessing an object caused subjects to rate it more positively in terms of attractiveness, value, and quality of design (but see Barone, Shimp, and Sprott 1997 for competing findings). Gibbs (1992) found that subjects who expected to sample a bitter-tasting solution twenty times rated the first taste as less aversive than did subjects who expected to taste it only once. Gibbs concluded that subjects who expected to consume the substance had manipulated their tastes for self-protective purposes to diminish the aversiveness of the event that they expected to be repeated so many times. The effect was relatively small, however, and an analogous effect was not obtained for a positive experience.

Several researchers have observed a positive relationship between the perceived likelihood of outcomes (manipulated experimentally) and the desirability of those outcomes, as would be predicted if people attempt to make the best of whatever outcome they expect to occur (Kay, Jimenez, and Jost forthcoming; McGuire 1960; Pyszczynski 1982). For example, on the eve of the 2000 presidential election, Kay and colleagues (forthcoming) found that supporters of both leading presidential candidates rated their preferred candidate more favorably if they were given information that suggested he was likely to win, and rated their dispreferred candidate less harshly if they were given information that made them pessimistic that he would win.

Temporal Proximity

A form of preference change that has received considerable attention in the economic literature is hyperbolic time discounting (see Angeletos et al., chap. 18 herein; O’Donoghue and Rabin, chap. 7 herein). Models of hyperbolic discounting have the property that people are much more impatient when it comes to trade-offs between immediate and slightly delayed gratifications than they are between delayed and slightly more delayed gratifications. Thus, for example, an individual might prefer one apple today to two apples tomorrow, but prefer two apples in a year and a day to one apple in a year (Thaler 1981). Hyperbolic discounting can produce systematic changes in preference (preference reversals).

Hyperbolic time discounting is somewhat different from the other determinants of preference change just discussed. Each of the other sources of preference change—habit formation, satiation, visceral factors, maturation, conditioning, social influences, and motivated taste change—leads to changes in hedonic experience at given points in time—that is, to changes in “instantaneous utilities.” Hyperbolic time discounting, though, is about changes in how a person trades off hedonic experiences at different times—that is, about changes in “inter-temporal utilities.”

Predicting Preferences

Knowing how people’s preferences change over time will go some way toward explaining and predicting their behavior. Yet many kinds of behavior cannot be understood without considering the extent to which people are able to predict the direction and magnitude of such change. To understand precommitment behavior, for example, we need to know not only what preferences the agent has at the time of the decision but also what preferences he or she expects to have at later points. Thus an adequate descriptive theory of choice cannot escape the question of how people predict future preferences.

Research on predicting preferences began with a 1990 paper by Kahneman and Snell. Subjects in that study ate ice cream and listened to music for ten days in a row and predicted how they would feel about the experience. Subjects’ preferences changed over time, although not in any obviously consistent fashion, and subjects predicted that their preferences would change, but the correlation between predicted and actual changes in preferences was close to zero.

Since publication of that study, research on predictions of future preferences has mushroomed. This section provides a broad overview of the literature, touching on what we view as some of the most robust and important findings. We organize the literature according to plausible underlying mechanisms that can lead to mispredictions. As should become evident from our discussion, note that many of the most robust phenomena may well be multiply determined.

Sources of Error in the Prediction of Future Preferences

Misconstrual One cause of people’s mispredictions is that they are mistaken about the objective qualities of the objects or events over-
which they have preferences. Wilson and Gilbert (forthcoming) refer to this phenomenon as the misconstrual problem. Clearly, if potential parents believe that having kids will be a continuously joyful experience, they will be disappointed when they find out that children get sick and keep their parents up all night. Similarly, people who buy lottery tickets may not know about the disadvantages to winning the lottery, and overlook the fact that they are likely to feel harassed by tax collectors, charities, and so on (Kaplan 1978). Misconstrual is likely to be a bigger problem when the decision concerns an unusual event. As Gilbert, Driver-Linn, and Wilson (forthcoming) note, “Many important events, such as marriage, the birth of a child, and terminal illness, are experienced just once or rarely, and thus we predictably mispredict how such novel events will unfold.” Since misconstrual is unlikely to lead to any systematic biases, however, we will have little to say about it further.

**Underprediction of Adaptation** The first studies to document a systematic bias in the prediction of future preferences were conducted by Loewenstein and Adler (1995). The studies differed from that of Kahneman and Snell (1990) in that they focused on preferences for an object as opposed to feelings about an experience. The studies examined whether people could anticipate changes in their preferences produced by the “endowment effect,” which is a very rapid form of taste change documented in the behavioral decision research literature (see, for example, Kahneman, Knetsch, and Thaler 1991). The endowment effect captures the observation that people become attached to objects in their possession and are reluctant to part with them, even if they had not particularly desired them in the first place. The endowment effect is an ideal form of preference change to examine because it occurs so rapidly—virtually instantly upon possession of an object.

Loewenstein and Adler (1995) report the results of two studies. In the first, subjects were shown a coffee mug, told that they were going to be given it in the near future, and that then they would have the opportunity to sell it back to the experimenter for cash. Subjects were asked to state the minimum price at which they thought they would be willing to do so. Subjects underestimated their own selling prices substantially, as if they failed to appreciate the fact that they would become attached to the object once they were endowed with it. In the second study, subjects were given an incentive for accurately predicting their own postendowment price. They were told that there was a 50 percent chance (based on a coin flip) that they would win a coffee mug, and that if they did win one, they would have a chance to exchange it for cash. Before flipping the coin, subjects stated a minimum selling price that would apply if they won the coin flip. Again, subjects who did not possess mugs substantially underestimated their own subsequent selling prices.

The most coherent explanation for the endowment effect is that it stems from adaptation and loss aversion—people adapt to ownership of the object, then are averse to losing it (Strahilevitz and Loewenstein 1998). Underappreciation of the endowment effect thus reflects underappreciation of adaptation to ownership. Subsequent studies suggest that underappreciation of adaptation applies to far more than ownership of objects. Thus, for example, Loewenstein and Frederick (1997) asked people to predict how much different types of environmental and personal changes would affect their well-being, and also to recall how much matched changes had affected their well-being in the past. Subjects expected future changes to affect their well-being substantially more than they reported that past changes had actually affected their well-being. The authors attributed this disparity in part to underappreciation of adaptation. Although the subjects had adapted to these changes in the past (which is why the changes had had little hedonic impact), they did not appreciate the degree to which the same process would occur in the future.

Gilbert and colleagues (1998) observed a similar pattern that they attributed to people’s underappreciation of their own psychological “immune systems.” In one study, they elicited assistant professors’ forecasts of how they would feel at various points after their tenure decision, and compared these forecasts to the self-reported well-being of others whose tenure decision had been made in the past. Current assistant professors predicted that they would be happier for a prolonged period following a positive decision and less happy after a negative decision, though they recognized that the long-term impact of the tenure decision would be minimal. In fact both groups converged to baseline levels of happiness shortly after the decision. In another study, Texas voters predicted how they would feel in the days after the gubernatorial election if their candidate (or the opposing candidate) won the election. Again, people overestimated the duration of both positive and negative feelings.

Also consistent with underappreciation of adaptation, Siefie, Dawes, and Loewenstein (1999) asked people who came to a clinic for an AIDS test to predict how they would feel approximately five weeks after obtaining a favorable or unfavorable test result. Respondents overestimated how good they would feel five weeks after obtaining a favorable result and (more tentatively, given the low rate of positive results) how bad they would feel five weeks after getting an unfavorable result.
Focalism  A second possible explanation for observed mispredictions of future tastes and feelings is what Schkade and Kahneman (1998) refer to as a focusing illusion and Wilson and colleagues (2000) call focalism. This is the tendency, as Loewenstein and Frederick (1997, 66) expressed it, to “overestimate the impact of any one factor on . . . quality of life . . . . Clearly, quality of life depends on a wide variety of different things, any one of which is likely to have only a small impact. However, perhaps when a respondent’s attention is focused on a particular type of change—e.g., in opportunities for fishing—they exaggerate its overall importance.”

The most compelling evidence for focalism comes from two sets of studies. In one, Schkade and Kahneman (1998) showed that midwestern college students believed students in California to be happier, and California students believed midwestern students to be less happy, despite the fact that both groups reported similar levels of happiness. The authors attributed this discrepancy to a focusing illusion based on the observation that those subjects who expected the greatest discrepancy were also those who, in subjective ratings of the importance of different aspects of life, put the greatest weight on weather as a determinant of happiness. In the other study, Wilson and colleagues (2000) had sports fans who watched a college basketball game predict how they would feel at various intervals after the game if the team they supported won or lost, then followed them up and measured their actual feelings in the ensuing days. Respondents overestimated how long they would feel bad if their team lost and how long they would feel good if their team won, but consistent with focalism this tendency was reduced when they were asked to think about other events in their lives.

Distinguishing Between Focalism and Underappreciation of Adaptation One potentially important instance of misprediction of future preferences is the discrepancy between patients’ and nonpatients’ evaluations of quality of life (QOL) associated with different medical conditions. In recent years, policy makers have sought to allocate scarce health-care resources in a rational fashion by taking account of how bad it is to have different conditions (and hence important it is to treat them). With the goal of providing inputs into such policy making, researchers have sought to measure the quality of life associated with different medical conditions. Repeatedly, however, it has been found that people who have various conditions judge their quality of life to be much higher than the general public anticipates their own quality of life would be if they had the same conditions.

Such discrepancies could result from either of the two causes of misprediction discussed here. Members of the public could be exaggerating the negative impact of medical conditions because their attention is being focused on them (a focusing illusion), or they could be underestimating their own ability to adapt to having such conditions. In a recent series of studies, Ubel and colleagues (2001) attempted to test these two explanations of the cause of the discrepancy between patients’ and nonpatients’ assessments of patients’ QOL. In some studies, they attempted a wide range of defocusing interventions, including one that was virtually identical to the one employed by Wilson and colleagues. For example, before being asked to estimate the quality of life with paraplegia, nonpatient respondents were asked to think of five aspects of their life that would change if they had paraplegia and five aspects that would not change. This and other interventions failed to raise nonpatients’ estimates of the quality of life of patients and rather tended to decrease them. In another study, Ubel and colleagues (2001) tested whether underprediction of adaptation could explain part of the discrepancy by having respondents think about times in their lives when they had suffered setbacks and about what had happened subsequently. Consistent with the idea that underprediction of adaptation is at least partially responsible for the discrepancy, thinking about their own experiences of adaptation indeed caused respondents to raise their QOL estimates significantly. While focalism appears to be important in some domains, it does not seem to contribute significantly to the discrepancy in QOL ratings between patients and nonpatients.

Hot-Cold Empathy Gaps At any point, one’s preferences depend not only on stable factors (such as long-lasting individual differences in tastes) but also on transient “visceral factors” (Loewenstein 1996). When visceral factors become elevated, they have a dual impact on preferences: they decrease one’s momentary well-being and they increase one’s preference for (and pleasure from) specific activities. Hunger, for example, is inherently miserable, and increases one’s immediate preference for food. Likewise, pain increases one’s desire for relief, sexual arousal for sex, and anger for aggression.

Again, when it comes to visceral factors, research has identified a systematic error in predicting changes in tastes (see Loewenstein 1996, 2000). When people are in a cold state—for example, not hungry, angry, sexually aroused, and so forth—they underappreciate what a hot state will feel like in the future and how such a state will affect their behavior. They make an analogous mistake when in a hot state and predicting how they will feel or behave when the heat dissipates (that is, when they are in a cold state). Such “hot-cold empathy
gaps" occur not only prospectively—when people predict their own past or future feelings and behavior—but also retrospectively, as in the infamous morning after syndrome, in which a cold morning self struggles to make sense of a hot past self's evening escapades (Loewenstein 1996; Loewenstein and Schkade 1999). Visceral states for which hot-cold empathy gaps have been documented include hunger (Read and van Leeuwen 1998), anxiety (Sieff, Dawes, and Loewenstein 1999), pain (Read and Loewenstein 1999), sexual arousal (Loewenstein, Nagin, and Paternoster 1997) and embarrassment (VanBoven et al. 2001). In the study of sexual arousal, for example, researchers found that male youths exposed to sexually arousing materials reported substantially higher likelihoods of behaving aggressively in a hypothetical date scenario than did youths not exposed to arousing material (Loewenstein, Nagin, and Paternoster 1997). Hot-cold empathy gaps also occur interpersonally: people have difficulty predicting the behavior of others who are in a different visceral state (VanBoven, Dunning, and Loewenstein 2000; VanBoven and Loewenstein forthcoming).

**Projection Bias**

Many of the findings in the literature on predictions of preferences, regardless of the mechanism postulated to cause them, appear to fit a simple pattern that Loewenstein, O'Donoghue, and Rabin (2001) label projection bias. People behave as if their future preferences will be more like their current preferences than they actually will be—as if they project their current preferences onto their future selves. Projection bias is analogous to a wide range of judgmental biases in which people's current state of knowledge contaminates their judgments of their own prior (or other persons') state of knowledge. Examples are the hindsight bias (Fischhoff 1975), in which people project their own current knowledge on themselves in the past (believing that they must have known in the past what they know now), and the curse of knowledge (Camerer, Loewenstein, and Weber 1989), in which, when people know something, they overestimate the likelihood that other people know it too.

In the paper introducing the concept of projection bias, Loewenstein, O'Donoghue, and Rabin (2001) model changing preferences as a matter of shifting state variables, as discussed earlier. Thus, for example, \( s \) could represent an individual's level of hunger and \( c \) their consumption of food, with

\[
\frac{\partial u(c,s)}{\partial c} > 0; \quad \frac{\partial u(c,s)}{\partial s} < 0; \quad \text{and} \quad \frac{\partial^2 u(c,s)}{\partial c \partial s} > 0.
\]

Projection bias says that when an individual attempts to predict the utility of future consumption \( c_t \) at a time when his or her actual state will be \( s_t \), that person's prediction will be biased by his or her current level of hunger, \( s_0 \):

\[
a_0(c_t,s_t/s_0) = (1 - \alpha)u(c_t,s_t) + \alpha u(c_t,s_0).\]

In words, the individual's predicted level of utility will lie between the actual level of utility he or she will experience in the future and the utility the person would experience with \( c_t \) given the current level of his or her state. People are then assumed to use these biased predictions to make decisions that affect the future, maximizing their biased prediction of their own future utility, instead of their true future utility.

In the case of hunger, this would mean that people who are not hungry who predict their preference for food at a point in the future when they can expect to be hungry would underestimate both their enjoyment of and desire for food at that time. Such a pattern is consistent with the folk wisdom that one should not shop on an empty stomach (see Gilbert, Gill, and Wilson 2002; Nisbett and Kanouse 1968) and has been demonstrated explicitly in research by Read and van Leeuwen (1998). In that study, office workers were asked to choose between healthy snacks and unhealthy snacks that they would receive in one week, either at a time when they could expect to be hungry (late in the afternoon) or satiated (immediately after lunch). Some were approached and asked to make this choice right after lunch, and some were asked to make it in the late afternoon. Not surprisingly, those who expected to receive the snack at a time when they were likely to be hungry were more likely to opt for the unhealthy snack, presumably reflecting an increased taste for unhealthy snacks in the hungry state. In addition, however, those who were hungry when they made the choice were also more likely to opt for unhealthy snacks than those who were satiated. People who were hungry when they made the decision seem to have anticipated being more hungry when they actually received the snack a week later—as if they projected their current hunger onto their future self.

The concept of projection bias can unify most of the misprediction results just discussed and many more; it can also make specific predictions in areas that have not yet been studied. Consider, for example, all of the different determinants of taste change (other than hyperbolic time preference and motivated taste change) discussed earlier. These comprise habit formation, satiation, refinement, visceral influences, maturation, and social influences.
Projection bias makes specific predictions for each of these forms of preference change. When it comes to habit formation, projection bias predicts that people will underestimate the impact of current consumption on utility from future consumption. Thus, for example, projection bias predicts that people will expect to enjoy increases in income for longer than they actually do. For satiation, the prediction is analogous. People will underpredict the size of their future of satiation, which will cause them to eat much too much of anything that they currently like—too much food at a restaurant, too many books or CDs by their favorite author or artist. Refinement is somewhat trickier, but projection bias again would imply that people undervalue its effects. Perhaps the best application of this phenomenon is interpersonal; according to this story, boors with crude tastes will tend not to appreciate the pleasures and nuances of experience conferred by refinement.

Turning to visceral influences, projection bias predicts that people who are in different visceral states—angry or not angry, hungry or not hungry, and so on—will have remarkably little empathy for or understanding of themselves in the different state. Much lack of empathy was demonstrated in a recent study by Van Boven and Loewenstein (forthcoming), in which people had just experienced (or were about to) a story about three hikers lost in the woods. Subjects were asked to write an essay about how the hikers felt, to predict whether the hikers would be more bothered by hunger or thirst, and to predict whether they themselves in that situation would be more bothered by hunger or thirst. The results were as predicted. Subjects who had experienced—and who could therefore be assumed to have been thirsty but not hungry—mentioned thirst earlier in their essay and predicted that they and the hikers would be more bothered by thirst than did those who had not yet exercised.

The pattern is very much the same for maturation. Projection bias implies that people will underestimate the extent to which their preferences will change as a result of aging, and that this bias will be most pronounced when tastes are most in flux. Many children, for example, find kissing scenes in movies disgusting and are unable to imagine that they would ever find kissing pleasurable to watch, much less do themselves. At the opposite extreme, one of the authors of this chapter observed that with age, his grandfather preferred sweeter wine, eliciting much amusement and contempt from his father. The author had the same reaction when his father’s taste in wine underwent the same shift at about the same age. (He is convinced of course that he will always prefer the dry wines that he currently enjoys.)

Projection also makes specific predictions about social influences.

First, people will underestimate the impact of future social influences (they will view their own tastes as more autonomous than they truly are). Second, they will exhibit a systematic bias when choosing whether to be a big fish in a small pond or a small fish in a big pond—they will tend to prefer bigger ponds than are actually optimal for their well-being (see Loewenstein et al. 2001). Imagine a faculty member at Podunk U. who is contemplating a move to Ivy U. According to projection bias, such a person will correctly imagine the reactions of his Podunk colleagues but will undervalue the impact that his reference group will quickly become the new colleagues at Ivy, who will be much less impressed with the fact that he is at Ivy.

Yet projection bias is not consistent with all of the preference-prediction biases documented in the literature. Phobics, for example (who are not currently experiencing fear of the thing they dread), do not project their fear onto their future self but exaggerate the fear they will experience if faced with the object of their phobia. While some teenagers underestimate the effects of maturation and falsely believe they will remain punk rock rebels for the rest of their lives, others overestimate the same effects and believe that they will adopt more mature tastes earlier than they actually do. In the latter case, people are aware that maturation is a determinant of taste change but exaggerate its effects. The “impact bias” of Gilbert, Driver-Linn, and Wilson (forthcoming) and Wilson and Gilbert (forthcoming) can account for such patterns. This bias is defined as a tendency to underestimate the enduring impact that future events will have on our emotional reactions. The relation between projection bias and impact bias is complex and the conditions under which we should expect projection bias or impact bias (or both) to be observed remains to be specified. One hopes that this will be explored in future research.

The model of the phenomenon of projection bias is uncertain in part because predictions of future preferences have only been studied for a narrow range of sources of preference change. No researchers to our knowledge have examined prediction accuracy for the types of changes in preferences caused by maturation or conditioning, and no research we know of has tested whether people’s attempts to refine their preferences are successful (let alone whether doing so actually makes people happier). Moreover, researchers studying predictions of taste change have not selected their domains at random. Indeed one could conjecture that researchers have focused on domains of behavior where they expected to observe bias (which is always more interesting than its absence). A useful role will be played by skeptical researchers who attempt to discredit (or provide evidence for the opposite) of the documented effects.
Honoring and Resisting Changes in Preferences

A theory of how people expect preferences to change over time will certainly help predict and understand their behavior. Such a theory, though, needs to be coupled with an account of when people choose to honor their current preferences, or, when they choose to honor preferences they expect to have when they experience the consequences of the decision. Take, for instance, a woman’s decision to eschew anesthesia during childbirth even though she expects to request it once labor begins. To understand such a pattern, one needs to explain why she wants to honor her immediate preferences rather than those she expects to have when the decision goes into effect. Thus a descriptively adequate theory of choice must also include an account of conditions under which people tend to go with current preferences and conditions under which they tend to go with future preferences. Similarly, a useful normative theory of choice would include an account of when we should honor current or future preferences.

On the descriptive side, existing literature reflects seemingly conflicting assumptions. As noted in the introduction, much of the literature on changing preferences consistently assumes that people aim to honor future preferences. This is particularly true of the literature on endogenous changes in tastes. Yet the literature on hyperbolic time discounting assumes throughout that people want to honor current rather than future preferences (for example, Laibson 1997). In the work of O’Donoghue and Rabin (2000), for example, it is assumed that, to the extent that people are sophisticated (that is, can predict the changes in their own future preferences wrought by hyperbolic time discounting), they will attempt to take actions that impose their current preferences on the future.

Of course a wide range of cases exist in which people attempt to satisfy future preferences. When grocery shopping on an empty stomach, we are often at least partly aware of the fact that we are likely to overestimate future consumption, and we decide to buy a little less than we would otherwise deem appropriate. When choosing an education, we think about how we will feel about the implicit career choice after graduation. To the extent that we do not take actions that satisfy our future preferences, often this is because we misestimate them, not because we do not want to satisfy them. At the same time, there are many situations in which we aim to honor current rather than future preferences.

This state of affairs raises the important question of when and why people sometimes aim to satisfy their current preferences rather than the preferences they expect to have when the consequences of the decision are experienced. Casual introspection points to two kinds of situation in which such a pattern occurs. The first (more common) situation occurs when people anticipate that they will not act in accord with their perceived self-interests, even though at some level they want to. This situation has received a tremendous amount of attention in the self-control literature, much of which focuses on strategies people use to ensure that their current preferences over future events will not be overridden by their future self (for example, Ainslie 1975; Hoch and Loewenstein 1991; Schelling 1984). The second situation involves cases in which one holds specific values with great conviction but worries that one will lose one’s current perspective—in effect becoming “corrupted” from the vantage point of one’s current perspective. This latter situation has received little if any attention in the empirical literature but has been of some interest to philosophers.

Why should people fear that they won’t act in their best self-interest? Perhaps the most common explanation of such failures involves the activation of visceral factors (Loewenstein 1996). Although visceral factors are essential mechanisms designed to produce specific behaviors (such as eating, drinking, fighting, flight, and copulation), when they become too intense they can override cognitive deliberations about self-interest. A wide range of self-destructive behaviors (for example, unsafe sex, eating problems, road rage, drug addictions, and so on) seem to stem from the activation of intense visceral factors. Most people are well aware of the power of visceral factors. Even at the moment of acting, people indeed may recognize that visceral factors are propelling them to behave self-destructively. Such awareness is, however, often insufficient to produce a correction in behavior, because self-control can require a prohibitive exertion of effort. Baumeister and colleagues (see Baumeister and Vohs, chap. 6 herein) have conducted numerous studies supporting the thesis that exercising self-control draws on some type of mental resource (sometimes referred to as willpower or ego strength) that is limited in quantity. Moreover, a wide range of situations—from alcohol consumption to sleep deprivation to immersion in a mob—seem to interfere with self-monitoring of behavior and hence with the exertion of self-control (Carver and Scheier 1998; Schelling 1984). In sum, if people expect future visceral factors to interfere with the pursuit of their own self-interest, we should not be surprised if they choose to reject such transient shifts in preference in favor of those that they currently hold.

Another situation in which people may favor their present over anticipated future preferences occurs when they are afraid of corrup-
tion—of abandoning values that they currently cherish. In some cases, people suspect that their preferences will change in ways that they find distasteful. Adolescents worry that they will become stodgy. Political radicals worry that they will become more conservative. Entering law students and medical students fear that the professional education they are about to receive will stamp out their idealism and focus them on the goal of making money. In these types of situations, one sometimes observes people taking efforts to commit to their current preferences.

We suspect that this kind of situation is much less common than the previous. Perhaps one reason is that when a preference is so strongly held that one would consider committing oneself to act on it in the future, to imagine that the preference will change over time is especially hard. Moreover, people rarely commit themselves to their current preferences in these situations. Perhaps we realize that dramatic action now—though in accord with current values—may make us unhappy in the future. While getting sterilized may be an effective means to prevent oneself from ever contributing to overpopulation, there is a significant probability that the procedure will lead to regretful misery later.

On the normative side, philosophers have worried about which preferences indeed should be honored in cases of conflict. The considerations proposed in this literature appear to provide rational underpinnings for the observed tendency to discount preferences due to visceral factors. The literature does not, however, unambiguously justify dismissing future preferences due to a general fear of corruption over time.

In the history of philosophy, David Hume is famous for arguing that preferences cannot be rationally criticized. In *A Treatise on Human Nature* (Hume 2000 [1739–1740], 267), he wrote,

‘Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger. ‘Tis not contrary to reason for me to choose my total ruin, to prevent the least uneasiness of an Indian or person wholly unknown to me. ‘Tis as little contrary to reason to prefer even my own acknowledg’d lesser good to my greater.

Even a follower of Hume, however, can argue that to act on preferences due to visceral factors may be unreasonable. Hume himself adds that a passion “can be call’d unreasonable” when it is “founded on false suppositions” (267). Suppose, for example, that a first drink induces the (false) belief that a second drink would make me more attractive, or increase my chances of befriending the person at the other end of the bar. Knowing that my preference for a second drink is based on a false supposition induced by the first, I may quite rationally dismiss that preference.

In more recent works, philosophers such as Richard Brandt seek to expand the range of desires that can be legitimately dismissed. Brandt (1998) provides a whole list of conditions under which preferences may legitimately be subjected to what he calls “rational criticism.” His view is that preferences are rational only if they would survive exposure to logic and the facts. Whether preferences can be rationally criticized, therefore, depends in part on whether they would go away if the subject clearly visualized the consequences of acting on them. Some preferences, Brandt argues, are due to a weakness of imagination, such as an inability to estimate the hedonic impact of alternative acts (Brandt 1998, 63). After playing tennis on a hot day, Brandt explains, a person may feel like drinking a couple of vodka and tonics; but if he made the effort to remember what happened the last time he succumbed to that particular desire, the desire would diminish dramatically (Brandt 1998, 67). Hence the preference can be rationally criticized. Also, in Brandt’s view, whether preferences can be rationally criticized depends on their causal history. *Authentic* preferences, on Brandt’s account, are acquired “from personal experience with liked/disliked examples of the target object” (1998, 71). Preferences that lack authenticity may be the result of “temporary emotional and motivational states,” of conditioning that took place under atypical environments such as a psychologist’s lab, or of imitation of someone mistakenly considered a proper role model (1998, 70–71). In Brandt’s view, then, to reject preferences due to visceral preferences if they are not “authentic” or would go away if certain consequences were made clear to the agent is perfectly legitimate.

Other considerations could potentially help us to determine which preferences deserve honoring in particular cases. Kusser (1998, 85) argues that one can legitimately discount “unhappy” preferences, by which she means preferences that are obstacles to happiness. Brandt (1998, 75) hints at this possibility: “For instance, you may have acquired an aversion to enjoying yourself, say by boating on a Sunday, because you have been taught to believe that a loving God wants the Sabbath day to be kept holy. Here a false or unjustified belief is the source of an attitude preventing happiness and desire-satisfaction.” Bykvist (1999, 51), like Brandt, suggests that the length of time that a desire is entertained matters: “There are cases to suggest that we should give more weight to long-lasting desires than to short ones.” If a newly acquired preference, perhaps expected to be fleeting, conflicts with a long-standing one, to honor the long-standing one is rationally
justified. These considerations too can be used to dismiss preferences due to visceral factors. If such preferences tend to be of brief duration, even fleeting, or if having them tends to make one unhappy—all of which seems plausible—they are open to rational criticism.

Yet philosophical considerations do not unambiguously support dismissing future preferences due to the mere possibility of corruption over time. As Derek Parfit (1986, 155) points out, while it may be true that my judgment will deteriorate over time, so that I “should give greater weight to what I used to value or believe,” it may also be true that my judgment will improve over time, as increasing knowledge and experience will help me make better decisions. Parfit (1986, 155) writes, “On this assumption, I should give to my future evaluative desires more weight than I give to my present evaluative desires, since my future desires will be better justified.” As Parfit notes, which of the two assumptions is better justified and under what conditions corruption is more likely than edification is unclear. For example, whether increasingly conservative politics results from a deeper understanding of the harsh realities of the world or from a decay in one’s moral and intellectual faculties remains an open question. So long as we do not know when to expect corruption or edification, the mere possibility of corruption does not give us grounds to rationally criticize future preferences.

Conclusions

In this chapter, we have discussed three interrelated issues: (1) how and why preferences change over time; (2) whether people can accurately predict how their preferences will change, and in what ways those predictions fail; and (3) when future tastes are expected to differ from current tastes, under what conditions people choose to indulge their current tastes or those they expect to prevail in the future. While discussion of these issues is relegated to separate sections of the paper, we have attempted to draw connections which are worth reiterating here in a more explicit fashion.

First, the errors people make in predicting their own preferences (issue 2) are likely to depend on the source of preference change (issue 1). To predict the effects of maturation, for example, people rely in part on intuitive theories of how tastes change over the life cycle. The accuracy of predictions will then depend on the correctness of the theories, as well as on the degree to which they apply to the person making the prediction. To predict one’s future appetite at a time when one is ravenously hungry, on the other hand, is likely to involve an attempt to remove from one’s prediction the effect of one’s immediate hunger. Thus, the processes invoked, and the types of errors one should expect to observe, depend on the nature of the underlying processes producing a change in preferences. Given this observation, it is perhaps somewhat surprising that many of the errors that have been observed in research on prediction of preference can be summarized by the simple pattern referred to as “projection bias.”

Second, whether people choose to indulge their immediate or their future preferences (issue 3) will again depend on the source of preference change (issue 1). As noted, people are much more likely to want to go with their future preferences when they reflect the influence of cumulated experience (for example, satiation or habit formation) than when they arise from the influence of intense but transient visceral factors. Our discussion of these issues was necessarily speculative because there has been remarkably little research on when and why people choose to indulge or deny anticipated preferences.

Third, decisions about whether to indulge or deny future preferences (issue 3), and the success of strategies designed to impose current preferences on the future, will depend on the accuracy of an individual’s predictions of future preferences (issue 2). Clearly, one can better take immediate measures to satisfy future preferences if one can predict what they will be. But predicting future preferences is also important when one wants to resist such change and satisfy one’s current preferences. One is much more likely to successfully practice safe sex in the heat of the moment, defy social pressures by “saying no” to drugs, or resist acting on angry feelings if one can predict when such feelings will arise.

A better understanding of changing preferences could shed light on the proper role of public policy. Economists typically assume that the correct role of government is to create and enforce rules under which people can maximally satisfy their desires (that are assumed to be fixed). This framework is inadequate if, as Sunstein (1993) notes in an essay titled “Democracy and Shifting Preferences,” preferences are not fixed but rather depend on public policies. “If the rules of allocation have preference-shaping effects,” Sunstein writes (1993, 202), “it is hard to see how a government might even attempt to take preferences ‘as given’ in any global sense, or as the basis of social choice. When preferences are a function of legal rules, the rules cannot be justified by reference to the preferences.”

The possibility that preferences may change as a result of public policy makes the Pareto criterion, as normally applied, inappropriate to evaluate government policies. Taking shifting preferences into account, we may in fact be able to support policies that would not oth-
erwise be justifiable. For example, consistent with the notion of refine-
ment introduced earlier, the government might want to institute poli-
cies intended to expose people to (or make it cheap or easy for them
to expose themselves to) things that they might not initially appreci-
ate. This could explain why, as Sunstein observes, people “support
nonentertainment broadcasting even though their own consumption
patterns favor situation comedies” (Sunstein 1993, 207).

Sunstein’s essay focuses not only on problems for public policy
caused by shifting preferences but also on those that result from fail-
ure to accurately predict or to honor future changes in preferences.
For example, regulation of drugs could be justified by the observa-
tion that people don’t seem to accurately predict the effect of addictive
drugs on their own subsequent preferences. Likewise (and analogous
to the idea that people do not always endorse anticipated changes in
their tastes), government might be concerned about not honoring or
creating in the first place certain types of preferences—for example,
“those that have resulted from unjust background conditions and that
will lead to human deprivation or misery” (Sunstein 1993, 203). Yet
government action only can be justified by reference to changing pre-
ferrences if the government itself can make accurate judgments about
such changes, something which—as this chapter has argued— is easi-
er said than done.

Governmental policy in the face of shifting preferences is likely to
confront some of the same issues that face decision makers them-
-selves—specifically, which preferences one should act on. The prob-
lem can be seen in the domain of health care policy, where predictable
changes in preferences abound. Whose preferences should we honor
when it comes to allocating scarce resources to preventing serious
chronic health conditions—that of patients or nonpatients, given that,
as we have already discussed, they differ systematically? In a differ-
ent health-related domain, healthy people often state that they are not
interested in receiving “heroic measures” if they become ill—chemo-
therapy, life-support systems, resuscitation—but sick people are far
more interested in receiving such care (Gleven et al. 1988). Whose pre-
ferrences should we act on? What if someone when healthy expresses a
desire to precommit to not receiving heroic measures? If such a desire
reflects a correct perception that powerful emotions evoked by sick-
ness will distort his or her decision making, then perhaps we should
allow the individual to precommit. Yet what if this is simply a hot-
cold empathy gap—a failure of someone in a cold state who can be
cavalier toward the prospect of death because he or she is not facing
it—to empathize with the true preferences the person will have once
he or she becomes sick?

Some economists have argued that economics should not be con-
cerned with issues of preference formation and change. Hayek, for
example, wrote, “If conscious action can be ‘explained,’ this is a task
for psychology but not for economics” (1948, 67). As Hodgson (1993,
154) writes in a discussion of this passage, “Hayek’s statement . . .
amounts to saying that we should not try to explain individual pre-
ferrences and purposes simply because such explanations are deemed
outside social science.” Becker (1976, 133) appears to concur with
Hayek’s judgment; he writes,

For economists to rest a large part of their theory of choice on differ-
ces in tastes is disturbing since they admittedly have no useful theory
of the formation of tastes, nor can they rely on a well-developed theory
of tastes from any other discipline in the social sciences, since none
exists. . . . The weakness in the received theory of choice, then, is the
extent to which it relies on differences in tastes to “explain” behavior
when it can neither explain how tastes are formed nor predict their
effects.

In Becker’s view, because of the lack of a theory of preference forma-
tion we should assume that tastes are stable over time and identical
across people.

We agree that no well-developed, useful theory of preference for-
formation and change exists as of yet, and there is remarkably little
rigorous data on which to construct such a theory. Yet we would ad-
vocate that economists respond to this not by leaving the task to psy-
cologists or by pretending that preferences remain stable but by
making the effort to study and model changes in preferences. A
deeper understanding of how preferences are formed and how they
change over time could increase the descriptive adequacy of eco-


This research was supported by NSF Grant SBR-9521914 to the Center
for Integrated Study of the Human Dimensions of Global Change at
Carnegie Mellon University, and by Jan Wallander and Tom Hedelius’
Foundation. We thank Kristof Bykvist, Shane Frederick, Niklas Karl-
sson, Ted O'Donoghue, Daniel Read, Paul Rozin, and Timothy D. Wilson
for helpful comments and suggestions.
Notes

1. Gilbert and Gill (2000, 394) argue that when we confront stimuli, we initially assume that our first reaction is veridical, then sometimes we elaborate and change our view. As they write,

[C]ognitive activities may normally be characterized by an initial moment of realism that is quickly followed by an idealist correction. According to this model, when people attempt to understand the objective properties of a stimulus, they automatically assume that their subjective experience of the stimulus is a faithful indicator of its properties, and then—if they have the time, energy, and ability—they rapidly “undo” that assumption by considering the possibility that extraneous factors may have shaped their experience.

2. Whether modeled with diminishing marginal utility or changing state variables, satiation is generally assumed to be a continuous process. That is, marginal increments to consumption are assumed to lead to, at most, marginal changes in the marginal utility of further consumption. Yet the assumption has rarely (if ever) been tested. Contrary to this assumption, our experience has been that satiation is discontinuous. We will listen to the same CD over and over with no discernible decline in liking until, with no warning, we reach a point where the idea of listening to it even one more time is distinctly aversive. Likewise, at various times we have become enamored of a particular kind of food and have indulged our new preference for it repeatedly but found that from one day to the next we go from loving the food to hating it. Such discontinuities in satiation, if verified by systematic empirical investigation, would help to explain why people find it so difficult to predict changes in their own tastes.

3. This problem may be possibly avoided by assigning low- and high-quality goods to separate mental categories. Zellner, Kern, and Parker (2002) gave experimental subjects samples of full-strength and diluted fruit juices. Diluted juices were rated worse by subjects who were told that all of the stimuli were “fruit juices” than by those who were told that the diluted fruit juices were “commercial drinks.” (No such effect of labeling was observed by subjects who rated only diluted juices.) In other studies reported in the same paper, consumers of premium coffee and beer who spontaneously assigned premium and low-grade items to separate mental categories rated the inferior-quality products more highly than those who did not.

4. See Carver and Scheier (1998) and Wilson, Gilbert, and Centerbar (forthcoming) for a further discussion of these issues.

5. Maturation is often difficult to distinguish from habit formation, since as one ages one often acquires experience with consumption in a fairly predictable fashion. For example, perhaps the general (though by no means universal) decrease in taste for candy as people enter adulthood results from satiation rather than biological changes associated with aging. Casual introspection suggests, however, that many of these changes are more tightly linked with age than they are to cumulative consumption experiences.

6. Yet in a seminal paper, Ben (1967) argues that many of the results that had been adduced in support of dissonance theory also could be explained in terms of what he called “self-perception theory.” Self-perception theory is based on the idea that people have imperfect knowledge of their own beliefs and tastes, and that they sometimes infer their own beliefs and tastes from their own actions, much as one would infer the beliefs and tastes of other persons by observing their actions. Self-perception theory would say that people increase their evaluations of things they choose, not to make themselves feel better about those things but because they naturally infer that they like the things they choose. Note that such a self-perception theory account of the taste-change phenomenon produced by dissonance researchers would not qualify as an instance of motivated taste change.

7. Gibbs’s finding may help to explain why experimental subjects who were made to expect to suffer (for example, by eating a worm or receiving an electric shock) were subsequently more likely to choose to suffer (by voluntarily eating the worm or taking the shock) than those who did not initially expect to suffer (for example, Aronson, Carlsmith, and Darley 1963; Comer and Laird 1975).

8. In the literature on predicting tastes, studies dealing with predictions of feelings are often discussed interchangeably with studies that focus on predictions of preferences. The two are certainly related—how much one wants something is likely to be related to how much one expects to enjoy it—but they are not synonymous (for example, Berridge and Robinson 1995).

9. The endowment effect is virtually instantaneous in the sense that people become attached immediately upon possession of an object. Yet this attachment intensifies over time as a function of how long the object is owned (Strahilevitz and Loewenstein 1998). Indeed one can think of the endowment effect as a special case of a more general influence of the history of object ownership on valuation. Valuation of an object increases as a function of whether one has owned it in the past, how long one owned it (either in the present or the past), and potentially (though not yet documented) as a function of how long ago one lost it if (one did) (Strahilevitz and Loewenstein 1998).

10. The study employed the Becker-Degroot-Marschak “truthful revelation” procedure that gives subjects an incentive for revealing their true valuations.

11. Wilson, Gilbert, and Centerbar (forthcoming) explain such phenomena in terms of sense making. They argue that humans have a strong tendency to make sense of events that actually occur, and that the process
lowers the intensity of emotional reactions. *Ordination neglect* is the term used to denote the failure to anticipate the effects of such sense making (Wilson and Gilbert forthcoming).

12. Visceral factors also can distort people's perception of self-interest through a process sometimes known as rationalization.

**References**


The Russell Sage Foundation

The Russell Sage Foundation, one of the oldest of America's general purpose foundations, was established in 1907 by Mrs. Margaret Olivia Sage for "the improvement of social and living conditions in the United States." The Foundation seeks to fulfill this mandate by fostering the development and dissemination of knowledge about the country's political, social, and economic problems. While the Foundation endeavors to assure the accuracy and objectivity of each book it publishes, the conclusions and interpretations in Russell Sage Foundation publications are those of the authors and not of the Foundation, its Trustees, or its staff. Publication by Russell Sage, therefore, does not imply Foundation endorsement.

BOARD OF TRUSTEES
Robert E. Denham, Chair
Alan S. Blinder
Christine K. Cassel
Thomas D. Cook
Larry V. Hedges
Jennifer L. Hochschild
Timothy A. Hultquist
Kathleen Hall Jamieson
Melvin Konner
Ellen Condliffe Lagemann
Cora B. Marrett
Eugene Smolensky
Eric Waner
Mary C. Waters

Library of Congress Cataloging-in-Publication Data
Time and decision: economic and psychological perspectives on intertemporal choice / George Loewenstein, Daniel Read, Roy Baumeister, editors.
P. cm.
"The papers were all first presented during a meeting at the Russell Sage Foundation in New York."—Intro.
Includes bibliographical references and index.
ISBN 0-87154-549-7
BF448 .T55 2003
153.8'3'—dc21
2002036743

Copyright © 2003 by Russell Sage Foundation. All rights reserved. Printed in the United States of America. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

Reproduction by the United States Government in whole or in part is permitted for any purpose.


Text design by Suzanne Nichols

RUSSELL SAGE FOUNDATION
112 East 64th Street, New York, New York 10021
10987654321