

**On the neural computation of utility: implications from studies of
brain stimulation reward**

Peter Shizgal
Center for Studies in Behavioral Neurobiology
Concordia University

In D. Kahneman, E. Diener & N. Schwarz, Eds. *Foundations of hedonic psychology: Scientific perspectives on enjoyment and suffering*, in press.

Acknowledgments

The author is grateful to Andreas Arvanitogiannis, Kent Berridge, Kent Conover, Randy Gallistel, Bart Hoebel, Daniel Kahneman, and Roy Wise for their helpful comments and to the “Fonds pour la Formation de Chercheurs et l’Aide à la Recherche du Québec” (grants ER-0124 and CE-0103), the Medical Research Council of Canada (grant MT-8037), and the Natural Sciences and Engineering Research Council of Canada (grant A0308) for research support.

Summary

1. Like other vertebrates, from goldfish to humans, rats will work in order to deliver electrical stimulation to certain brain sites. Although the stimulation produces no evident physiological benefit, it is sought out avidly, as if it were a biologically significant resource. Thus, it has long been thought that the rewarding stimulation activates neural circuitry involved in the evaluation and selection of goals.
2. Computing the utility of goal objects involves a tightly integrated set of perceptual, cognitive, and motivational mechanisms. I argue that rewarding electrical brain stimulation engages only a subset of these mechanisms. If so, comparison of the ways in which the utility of electrical brain stimulation and natural reinforcers are computed may highlight operating principles and isolate components of the computational mechanisms.
3. In the view proposed here, information about goal objects and consummatory acts is processed, in parallel, in three different channels.
 - 3.1. Perceptual processing indicates what and where the goal object is
 - 3.2. A stopwatch-like interval timer predicts when or how often the goal object will be available.
 - 3.3. Under the influence of information about the current physiological state, an evaluative channel returns a subjective weighting of strength variables such as the concentration of a sucrose solution or the temperature of an air current.
 - 3.4. The output of these channels is recorded in multidimensional records that include
 - 3.4.1. information of perceptual origin about amount and kind (e.g., food, water, or salt)
 - 3.4.2. information from the timer about rate and delay
 - 3.4.3. a subjective assessment of intensity provided by the evaluative channel

4. This chapter addresses the relationships between brain stimulation reward (BSR), the perceptual, interval timing, and evaluative channels, and the variants of utility proposed by Kahneman and his coworkers on the basis of their studies of evaluation and choice in human subjects.
 - 4.1. It is argued that the output of the evaluative channel can be manifested in experience as pleasure or suffering but that awareness is not necessary in order for this signal to influence action.
5. The neural signal injected by rewarding electrical stimulation is portrayed as providing meaningful information about rate, delay and intensity but not about amount or kind. This proposal is used to account for
 - 5.1. competition and summation between BSR and natural rewards
 - 5.2. differential effects of physiological feedback on the utility of BSR and natural rewards
 - 5.3. matching of behavioral allocation to the relative rates and intensities of BSR
 - 5.4. differences in the elasticity of demand for BSR and food in a closed economy
 - 5.5. the high substitutability of BSR for food and water in an open economy
6. The powerful aftereffect of BSR that potentiates efforts to obtain additional stimulation is related to expectancy.

Brain stimulation reward

A rat sits quietly in the start box of a runway, its access to the 6-ft alley blocked by a acrylic panel. The rat begins to groom, licking its paws and rubbing them over its snout. As the fastidious creature cleans the top of its head, it encounters a miniature electrical connector fastened firmly to its skull; the rat's paws sweep across this now-familiar appendage without breaking rhythm. Unhindered by the flexible cable linking the connector to a stimulator, the rat then turns its head to groom its flank.

The grooming bout is cut short as the stimulator sends small surges of current through an electrode attached to the connector. Deep in the brain, these stimulation pulses trigger salvos of impulses in nerve fibers coursing past the electrode tip. The rat looks up, highly alert. It stretches forward and explores the start box, sniffing the floor and scanning its head from side to side. At the offset of the stimulation, the rat approaches the acrylic panel, seizes the top with its forepaws, and hops as if trying to vault the barrier. Shortly thereafter, a solenoid withdraws the panel, and the rat explodes into the alley like a sprinter quick out of the blocks. It races to the goal box, and without breaking stride, presses a lever that has triggered delivery of stimulation pulses on previous trials.

On this trial, the rat is disappointed: The experimenter has turned off the stimulation in the goal box. Accustomed to such betrayals, the rat adjusts to the new conditions within a few trials. Although the stimulation in the start box is unchanged, it now fails to galvanize the rat's behavior. No longer expecting to receive stimulation in the goal box, the rat remains in the start box following removal of the barrier, and after casually sniffing the slot into which the barrier has been withdrawn, the rat lies down and yawns.

In this vignette, the rat treats the stimulation much as if it were a biologically significant resource, such as food during a period of limited availability or a heated nest box in a cold environment. Given that such natural resources serve as highly effective inducements for learning and performance, the effect of the stimulation that the rat seeks to reinstate is called "brain stimulation reward" (BSR).

The phenomenon of BSR has been observed in a wide variety of vertebrates, from goldfish to humans (Bishop, Elder, & Heath, 1963; Boyd & Gardiner, 1962; Distel, 1978; Lilly & Miller, 1962; Olds & Milner, 1954; Porter, Conrad, & Brady, 1959; Roberts, 1958). Unlike the case of natural reinforcers, the rewarding effect of electrical brain stimulation is not undermined by recent “consumption,” and no prior deprivation of essential physiological resources is required in order to elicit and maintain vigorous performance. Among the feats that rats will perform to obtain the rewarding stimulation are running uphill while leaping over hurdles (Edmonds & Gallistel, 1974) and crossing an electrified grid (Olds, 1958). When brief, intense trains of lateral hypothalamic (LH) stimulation are available continuously, rats will work for the rewarding stimulation at the expense of forgoing their sole daily opportunity to eat (Frank & Stutz, 1984; Routtenberg & Lindy, 1965) and will prefer such stimulation to water, even following prolonged fluid deprivation (Morgan & Mogenson, 1966). Many abused drugs, such as cocaine, amphetamine, heroin, cannabis, and nicotine, potentiate the rewarding effect of the stimulation (Wise, 1996).

The rewarding effect of electrical brain stimulation has long been linked to the subject of this volume: the scientific study of enjoyment and suffering. Indeed, the first report of BSR in the press (Macfarlane, 1954) heralded the discovery of a “pleasure area” in the brain, and similar phrases were used in early scientific publications (Olds, 1956). However, subsequent developments, both in the study of BSR and in the study of the relationship between hedonic experience and choice, argue for a richer and more nuanced characterization.

Toward a new view of brain stimulation reward

In this chapter, I summarize and extend a new view of the neural signal mimicked by the rewarding stimulation (Shizgal, 1997; Shizgal & Conover, 1996). This view is rooted in a long-standing idea: that multiple modes of processing are brought into play in parallel when an animal encounters a goal object (e.g.(Pfaffmann, Norgren, & Grill, 1977; Zajonc, 1980)). Perceptual processing determines the identity, location, and physical properties of the goal object, whereas the information derived from evaluative processing is used to determine what the goal object is worth. A third processor, which acts as a stopwatch timer, is concerned with when or how often

the goal object will be available (Gibbon, 1977; Gibbon, Church, Fairhurst, & Kacelnik, 1988). Kent Conover and I have proposed that BSR arises from activation of the evaluative system (Shizgal, 1997; Shizgal & Conover, 1996). In our view, the stimulation simulates a meaningful signal in this system and also provides an interpretable input to the timer. The evaluative and timing systems provide sufficient information for computing a payoff. Thus, the rat is drawn back to the site where stimulation was delivered previously and incited to repeat the acts that trigger the stimulation. In contrast, the stimulation-induced signals are not interpretable by the perceptual system. If so, the rewarding stimulation cannot recreate the perceptual experience produced by contact with a natural goal object. In effect, the rat knows that the lever delivers something valuable, but it cannot determine what that something is.

The account developed here embeds ideas that Conover and I derived from studying BSR and gustatory reward in rats within a broader conception of how utility is computed and links these ideas to concepts drawn from the study of hedonic experience in humans. From this perspective, there are several ways in which the analysis of BSR might advance the scientific study of enjoyment: 1) The natural stimuli that give rise to enjoyment act on the perceptual, evaluative, and timing systems. If rewarding brain stimulation indeed mimics only a subset of these effects, then it can serve as a tool to tease apart what is normally a tightly-integrated complex of responses, helping isolate individual components so that their properties can be described and understood. 2) In the view presented here, the neural signals mimicked by BSR bias the subject to resist interruption of pleasurable activities and to repeat actions that have led to pleasant consequences in the past. If so, a thorough understanding of the processes underlying BSR should help explain how the evaluative system steers ongoing behavior and influences both the amount and kinds of enjoyment that will be experienced in the future. 3) The neural signals underlying BSR arise from an observable volley of nerve impulses elicited at a known location in the central nervous system. Thus, the phenomenon of BSR should be particularly propitious for attempts to understand evaluative processing in terms of the activity of identified neural circuitry. (For reviews of research aimed at identifying the neural circuitry subserving BSR, see (Shizgal, 1997; Shizgal & Murray, 1989; Yeomans, 1988).)

Variants of utility

In recent years, Kahneman and his coworkers have investigated the relationship between choice and the operation of the evaluative system, delineating several variants of utility (Kahneman, 1994; Kahneman, Wakker, & Sarin, 1997). Their work has been carried out with human subjects, whereas research on BSR has been conducted almost exclusively on rats and other laboratory animals. Nonetheless, the distinctions drawn by Kahneman et al. between different types of utility and their proposals as to how variants of utility are interrelated provide a very useful framework for linking research on BSR to more general conceptions of choice and to hedonic experience.

Instantaneous utility. We continuously evaluate the stream of sensory experience and adjust our behavior accordingly. Kahneman et al. (1997) refer to the product of such ongoing evaluation as “instant utility,” a quantity that can vary in sign and magnitude. They give the term two connotations, one defined with respect to hedonic response and the other with respect to action. According to this view, instant utility is experienced along an opponent hedonic dimension (“good/bad”) while biasing the individual to continue or terminate the current course of action. States and stimuli that produce positive values of instant utility are experienced as pleasurable while impelling us to continue what we are doing; states and stimuli that produce negative values have the opposite effects. Like the brightness of a visual stimulus at a particular point in time (Schreiber & Kahneman, submitted), instant utility is a property of the moment. I shall refer to this quantity as “instantaneous utility.”

Remembered utility. During experiences such a meal at a fine restaurant or a visit to the theater, instantaneous utility fluctuates over time. Despite the complexity of the resulting temporal profiles, we have little difficulty in applying single ratings to such experiences such as “four stars out of a possible five” or in reporting to others whether we obtained our money’s worth. Kahneman et al., refer to such unitary ratings of temporally extended experiences as “remembered utility.” They propose that in compressing a temporal profile of instantaneous utility into a

remembered utility, we apply heuristics that simplify the task, speed its execution, and minimize the mnemonic and computational resources required.

Decision utility. Remembered utility influences behavior via a further computation, the calculation of the weights applied to the different outcomes of a decision under active consideration. Kahneman et al. call these weights “decision utilities” (Kahneman, 1994; Kahneman et al., 1997). In their terms, the rat’s choice of whether or not to leave the start box would be said to depend on the decision utilities of two outcomes, obtaining additional stimulation in the goal box or lying down and resting in place.

Predicted utility. Kahneman et al. point out that the calculation of decision utility may reflect not only remembered utilities but also additional factors, such as the “predicted utility” of the outcome. Whereas remembered utility returns the overall “goodness” of the last meal we ate at a particular restaurant, predicted utility reflects our expectation of how much we will enjoy a visit to that restaurant today.

Relationship of BSR to different variants of utility

How is the ongoing neural activity driven by rewarding electrical stimulation related to instantaneous utility? According to the proposal advanced by Conover and Shizgal, the rewarding stimulation achieves its grip over ongoing behavior by simulating the real-time effect of a natural reward on the evaluative system, i.e., by driving instantaneous utility to positive values. I propose that this signal can steer behavior in the absence of awareness but does not do so invariably. Via the allocation of attentional resources, the instantaneous utility signal can gain access to working memory and may be manifested in human experience as pleasure or suffering. Thus, the dual meaning imparted to instantaneous utility by Kahneman and his coworkers has been retained, but the link between the action component and the hedonic component is weakened, with the action component treated as the more fundamental. Some advantages of allowing the instantaneous utility signal to impinge on awareness are discussed below.

In the vignette at the beginning of this chapter, the behavior of the rat depends on whether or not it has received sufficiently rewarding stimulation in the goal box on preceding trials. Thus,

the rat appears to have recorded the utility of the stimulation received previously. We will see shortly that there is a striking similarity in the way that the instantaneous utilities of BSR in rats and certain temporally extended experiences in humans are translated into remembered utilities.

Records of payoff are inherently multidimensional and may well be derived from multiple modes of processing. I argue below that what Kahneman et al. call remembered utility captures the subjective “intensity” of the reinforcer, which is but one of several components of such records. The evaluative channel that assesses intensity is complemented by a stopwatch timer that delivers assessments of encounter rate and delay and by perceptual mechanisms that can return estimates of amount (e.g., the mass of an acorn) and kind (e.g. food versus water). Information about kind can be used to determine the degree to which one reinforcer can substitute for another. A key tenet of behavioral economics is that substitutability determines whether and how much behavioral allocation will shift from one reinforcer to another in the face of price changes. Moreover, it can be argued that the elasticity of demand for a particular kind of resource depends on additional information of perceptual origin: environmental distribution of that resource in the environment.

In this view, decision utilities are derived from a combination of perceptual, timing, and evaluative data. If BSR indeed reflects meaningful signals in the evaluative and timing systems in the absence of meaningful perceptual information, then performance for BSR should respond differently to economic constraints than performance for natural reinforcers. The literature reviewed below is interpreted to support this contention and to suggest that comparisons between performance for BSR and for natural reinforcers shed light on the psychological resources involved in computing decision utilities.

In the opening vignette, the start-box stimulation produces an aftereffect that potentiates behavior aimed at procuring additional reward. If the rat has been given a free “taste” of the stimulation at the start of the trial, it will show more pronounced anticipatory behaviors prior to the removal of the barrier and will run down the alley faster once allowed to enter. This is reminiscent of the way that savoring a particularly tasty hors-d’oeuvre at a reception can incite

visual search for the waiter and vigorous pursuit once he reappears. Just as the anticipatory search and the pursuit of the waiter depend on the expectation that the supply of hors-d'oeuvres has not yet been exhausted, the anticipatory behavior of the rat in the start box depends on the expectation that stimulation will be available in the goal box. Such expectations may be related to the predicted utilities discussed by Kahneman et al. in that anticipation of a future event influences present choices.

In order to be manifested in behavior, a decision utility must be processed by a selection rule. "Choose the largest," will be assumed as the rule. The question of how the resulting decisions are translated into action is beyond the scope of this chapter; the reader is referred to Gallistel's *Organization of Action* (1980) for a fine introduction to this topic.

In the following sections, the relationship between BSR and the variants of utility proposed by Kahneman et al. is discussed in detail. Exploration of this relationship casts BSR data in a new light and suggests new directions for future research.

BSR basics

Some basic characteristics of the electrical stimulus and the neural circuitry responsible for its rewarding effect must be described before developing the arguments linking BSR to different variants of utility. The stimuli used in most modern BSR experiments consist of trains of short-duration current pulses. With pulse duration held constant, the strength of a train is determined by pulse amplitude (current) and frequency. The greater the current, the larger the number of neurons directly stimulated by the electrode. Over the ranges of frequencies used in most studies, each directly stimulated neuron can be assumed to fire once per pulse. Thus, as depicted in Figure 1, each pulse produces a synchronous volley of nerve impulses (action potentials) in the population of directly-stimulated cells that give rise to the rewarding effect, and the aggregate firing rate of this population is determined by the stimulation current and frequency. It is highly unlikely that this population of neurons responds in such a synchronous manner to any natural stimulus, yet the artificial stimulation does mimic some of the properties of a natural reinforcer.

As discussed below, this provides a clue as to how information is represented in the neural system underlying the rewarding effect.

The counter model. The firings of the directly stimulated neurons appear to be translated into the rewarding effect in a surprisingly simple manner. With the duration of a stimulation train held constant, the strength of the rewarding effect appears to depend only on the aggregate rate of firing in this population of directly stimulated cells. According to this “counter model” (Gallistel, 1978; Gallistel, Shizgal, & Yeomans, 1981; Simmons & Gallistel, 1994), it matters not whether one hundred directly stimulated neurons fire ten times each during a particular time window or whether twenty neurons fire fifty times each. The rewarding impact of the stimulation will be the same provided that aggregate impulse flow is constant. If activity elicited in these neurons by natural stimuli is integrated in the same manner, then the synchronous firings triggered by the artificial stimulation should produce the same effect as an equivalent number of asynchronous firings triggered by a natural stimulus.

Figure 1 about here

The counter model is shown in Figure 1. The directly stimulated neurons responsible for the rewarding effect are depicted as providing input to an “integrator” that combines the effects of incoming action potentials over time and space. The output of the integrator is determined by the aggregate rate of firing at its input. It is argued here that this output is the sole determinant of the instantaneous utility of the stimulation and that the remembered utility of the stimulation is derived from certain “exemplar values” (Schreiber & Kahneman, submitted) of its instantaneous utility.

Instantaneous utility, resistance to interruption, and BSR

At many brain sites, BSR is accompanied by aversive side effects, which are due to the activation of different neurons than the ones responsible for the rewarding effect (Bielajew & Shizgal, 1980; Bower & Miller, 1958; Shizgal & Matthews, 1977). By judicious selection of stimulation site, current, frequency, and temporal pattern, the aversive side-effects can be minimized. When such precautions are taken, the rat will readily press a lever to initiate a long-

duration train of stimulation but will not press a second lever that turns off the stimulation (Shizgal & Matthews, 1977). If the experimenter interrupts such a train, the rat will immediately rush over to the lever and reinitiate the stimulation. This suggests that if given the opportunity, the rat would strive to prevent interruption of such stimulation. If so, it should prove possible to measure the action component of instantaneous utility in real time by assessing the commitment of the rat to keeping the current flowing.

To my knowledge, such an experiment has not been done. A promising way to perform it would be to use the temptation of an alternative reward to assess the rat's commitment to the ongoing stimulation. At different times during the delivery of a long train, a choice would be offered between two options: a) continuation of the train or b) immediate cessation of the train coupled with delivery of an alternate reward. The strength of the alternate reward required to tempt the rat to terminate the long train would provide a measure of the instantaneous utility of the ongoing stimulation at the moment of choice. In principle, such a measure would provide a direct test of the prediction that both the remembered utility of the stimulation and its instantaneous utility are derived from the output of one and the same integrator.

Instantaneous utility, hedonic experience, and BSR

We experience pleasure and pain as powerful, adaptive influences on our behavior. As Bentham put it so memorably (Bentham, 1996 (originally published: 1789)):

“Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point out what we ought to do, as well as to determine what we shall do.”

Bentham's assertion that pleasure and pain *direct* action may ring so true to our experience as to lead us to the converse view: that action *reflects* hedonic state. If pleasure and pain lead us to seek out or maintain contact with a pleasurable stimulus and to avoid or interrupt contact with a painful one, then is it not justified to infer these experiences from observation of the acts they promote? It is in this sense that the description of BSR sites as “pleasure areas” has had an

intuitive appeal to many. If the rat is willing to work so hard to initiate the electrical stimulation, then must not the stimulation be pleasurable?

In my view, the answer to both questions is no, not necessarily. I will argue that the two components of instantaneous utility can coincide, but that they need not do so invariably. Thus, the link between hedonic experience and the control of action is less direct in the account presented here than in Bentham's original formulation or in modern developments of Bentham's position (Cabanac, 1992). I speculate below on what may be gained by supplementing the action-oriented component of instantaneous utility with a hedonic response.

To turn Bentham's formulation around and to infer pleasure and pain from behavior is to make a strong assumption: that actions such as resistance to interruption of a stimulus or attempts to escape from it cannot be produced in the absence of a hedonic response, i.e., without the express consent of the "sovereign." ("It is for them *alone* ... to determine what we shall do.") By labeling a state as pleasurable or painful, we imply that we are aware of it; "unconscious pleasure" and "unconscious suffering" are oxymorons. If so, asserting that a hedonic response is a necessary condition for resistance to interruption or escape is tantamount to stating that such actions cannot be produced in the absence of awareness.

Bentham's position does not stand up well in the face of a large body of psychological research and theory that treats much of the foundation of perception, thought, emotion and action as hidden from awareness (Baars, 1988; LeDoux, 1996; Nisbett & Wilson, 1977). In such views, consciousness depends on the serial operation of a limited-capacity process. Rather than forcing all signals vying for the control of action to pass through this processing bottleneck, much of the task of real-time control is assigned to a collection of specialized lower-level processors operating in parallel and in the absence of awareness (Baars, 1988). If so, the hedonic and action-oriented components of instantaneous utility are dissociable, and we cannot necessarily infer the hedonic content of experience on the basis of behavioral observation alone.

In human subjects, we can address the relationship between hedonic experience and the control of action empirically, via methods for concurrent measurement of self-ratings and

behavior. For example, ratings of the sign and intensity of hedonic experience can be collected while observing whether the subject maintains or breaks off contact with a stimulus. It is not surprising that strong correlations have been noted in such studies between subjective hedonic ratings and measures of choice (Cabanac & LeBlanc, 1983). However, dissociations have been noted as well. For example, in a study of heroin addicts self-administering morphine, low doses of the drug were vigorously self-administered despite subjective ratings of zero on both a monetary value scale and a Likert scale of liking; a saline solution received similar subjective ratings but was not self-administered (Lamb et al., 1991). It is interesting to note that increasing the dose of morphine brought the subjective ratings into accord with the behavioral measure. Thus, self-ratings of hedonic response could coincide with action but did not do so invariably. (The reader is referred to the chapter in this volume by Berridge for an alternative interpretation of these data.)

In contrast to the tools available for studying the relationship between hedonic experience and action in humans, we do not have well-validated and general means for measuring enjoyment and suffering in non-human animals. Although it should prove possible, as proposed above, to measure the rat's resistance to interruption of the stimulation, we cannot be sure how the rat feels while the current is flowing. Nonetheless, I will propose that the relationship between the control of ongoing behavior and processes that contribute to awareness in humans could be investigated by neurobiological means in non-human subjects.

In the view elaborated here, which borrows from proposals by Ledoux (1996), the output of the neural process that determines whether an action will be continued or terminated (the "continue/stop signal:" the action-oriented component of instantaneous utility) is not isomorphic with pleasure or suffering but will be manifested in awareness as a hedonic response if the continue/stop signal gains access to working memory. This access is gated by attention and will be most likely to occur when the action-oriented signal and associated stimuli attain high values. Nonetheless, sufficient allocation of attention might allow weaker signals to trigger a hedonic response, and strong signals might fail to do so in the absence of attention, e.g., when events are highly predictable and behavioral responses are highly practiced. When the continue/stop signal

does succeed in breaching the waterline of awareness, it can marshal further attentional resources and direct planning while coordinating the activity of processes that operate beyond the margins of conscious experience.

To develop the argument, let us assume that we were given the task of designing a robot that simulates the behavior of a rat. Ongoing action is controlled by a continue/stop signal derived from real-time information about external stimuli and from the state of the internal environment. For example, when body temperature is low and a warm microenvironment is encountered, the continue/stop signal will have a positive value, thus promoting continued contact with the heat source and a return to thermal homeostasis; when internal temperature is too high, the same thermal stimulus will drive the continue/stop signal to negative values and terminate contact. This adjustment of the neutral point as a function of internal state reflects Cabanac's concept of alliesthesia (Cabanac, 1971).

The sensory and evaluative mechanisms that generate the continue/stop signal should allow the robot to simulate certain adaptive responses of a rat to ongoing sensory stimulation. However, the robot would need additional circuitry in order to mimic abilities that figure prominently in cognitively-oriented accounts of goal-directed behavior such as the resolution of conflicts between multiple goal-related stimuli by means of selective allocation of attention, navigation in space using stored representations, and planning a route leading from the current state to a higher-valued one. In an influential treatment of animal navigation (Gallistel, 1990), an egocentric spatial representation of the current environment is constructed from successively encountered stimuli and then translated into geocentric coordinates using stored information about the position of vantage points and angles of view. Essential to such tasks is a readily accessible ("working") memory store, in which critical information is held on-line. Working memory also plays an essential role in models that find efficient routes to goals (Gallistel, 1990; Johnson-Laird, 1988; Miller, Galanter, & Pribram, 1960); access to this limited-capacity store is gated by attention. If we were to incorporate attention, working memory, and a route-finding

mechanism in our robot, would it be advantageous to allow the continue/stop signal to interact with them, and if so, how?

For help in addressing this question, let us enlist the assistance of a wise student of behavior, the novelist, Joseph Heller. In *Catch-22* (Heller, 1961), Heller explores the relationship between instantaneous utility and the control of action. Heller's antihero, Yossarian, is a World War II bombardier surrounded by suffering, death, and destruction. Desperate to survive his military service, Yossarian refuses to accept that things are as they ought to be. During a philosophical argument about the failings of the Supreme Being, Yossarian complains:

"Why in the world did He ever create pain? ... Why couldn't He have used a doorbell instead to notify us, or one of his celestial choirs? Or a system of blue-and-red neon tubes right in the middle of each person's forehead. Any jukebox manufacturer worth his salt could have done that. Why couldn't He?"

Yossarian wants to believe that the protective function of pain could be fulfilled by a warning signal that is merely informative. But what would confer upon such a signal the ability to capture attention, wrest control of planning, and coordinate the multiple processes controlling action? Sadly for Yossarian and his flak-riddled comrades, the insistent unpleasantness of pain is a highly effective means of achieving these ends. Gladly for the rest of us, so too is the shock of intense pleasure.

With our response to Yossarian in mind, let us return to our robotic rat and implement a key improvement. We will now arrange things so the continue/stop signal can access working memory and influence the allocation of attention. The continue/stop signal will attain large negative values in response to ongoing or imminent tissue damage; focusing attention on such a signal would tend to promote it to the top of the planning agenda and reduce the odds that competing stimuli would divert scarce cognitive resources from the task of terminating the noxious input. In contrast, the continue/stop signal will register large positive values in response to contact with potentially beneficial stimuli such as food sources or locations that promote

maintenance of thermal neutrality. The likelihood of interrupting contact with the beneficial input would be reduced by allowing it to draw attention away from competing stimuli. Temporarily suppressing planning might also help “lock in” contact with an input that is driving the continue/stop signal to large positive values.

Working memory and attentional control of its input are regarded as key components of the foundation for awareness in humans (Baars, 1988; Johnson-Laird, 1988; LeDoux, 1996). Thus, if this sketch were generalized from our robotic rat to ourselves, we should predict that extremes of instantaneous utility (strong continue/stop signals) would tend to be reflected in our experience as well as in our behavior. Stimuli that produce weaker excursions of instantaneous utility might, nonetheless, exercise behavioral control, but, as in the case of the low doses of morphine self-administered by the addicts, such stimuli are less likely to be manifested in awareness.

It has been argued that consciousness enables the “broadcasting” of information throughout the cognitive architecture to the many specialized processors that operate beyond the margins of awareness (Baars, 1988). If so, expressing the continue/stop signal in awareness as pleasure or pain would help marshal and coordinate the activity of multiple cognitive processes in mounting a highly integrated response to the eliciting stimuli.

I propose below that direct electrical stimulation of certain brain regions can mimic the effect of a naturally occurring stimulus on the neural circuitry that computes instantaneous utility. If so, the argument developed above predicts that such stimulation should be able to drive instantaneous utility to levels that can impinge on awareness in humans. Indeed, when direct electrical stimulation has been delivered to some of the brain regions homologous to sites where BSR is obtained in rats, its effect has been described by human subjects as pleasurable (Heath, 1964).

By means of electrophysiological recordings, the activity of neurons implicated in working memory is monitored routinely in non-human animals (Goldman-Rakic, 1996; Watanabe, 1996), and the modulating effects of attention can be observed (Treue & Maunsell, 1996). Thus, it may

prove possible to determine, by conventional neurobiological means, whether instantaneous utility signals can capture attentional resources and gain access to working memory. Research on BSR could play a crucial role in such experiments by identifying neural circuitry subserving instantaneous utility and by providing a potent means of controlling it.

Instantaneous utility is a property of the moment, and thus, this discussion had been focused on real-time processing. Let us now turn to a process that bridges the present and the future: translation of instantaneous utility into a stored record. The bulk of the research carried out on BSR has probed such records.

Transformation of instantaneous utility into remembered utility

In Kahneman's portrayal, remembered utility is derived from a temporal profile of instantaneous utility. The continuously fluctuating value of instantaneous utility during a temporally extended experience is compressed into a single remembered utility on which future decision weights can be based. Imagine, for example, that upon pressing the lever in the goal box, the rat receives a prolonged train of stimulation that waxes and wanes in strength over several minutes (e.g. (Lepore & Franklin, 1992)), much as the level of a drug in the bloodstream rises and falls following administration. How does the rat compress this temporally extended experience over time so as to derive a single decision utility that can weight future choices?

One way to derive a single remembered utility from a temporally extended experience is to compute the temporal integral or the average of the entire sequence of instantaneous utilities. Kahneman and his coworkers propose a very different and much simpler strategy. Their human subjects appear to extract two key values from the temporal profile: the peak instantaneous utility and the instantaneous utility at the end of the experience; some intermediate value, such as the average of the peak and end values, is then used as the remembered utility (Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993; Redelmeier & Kahneman, 1996). Only in unusual circumstances would the peak-and-end rule be expected to generate a result radically different from the outcome of temporal integration. However, a simple rule of thumb, such as peak-and-

end averaging, should be executed more quickly than retrospective temporal integration, while consuming fewer mnemonic and computational resources.

If the peak-and-end heuristic were employed, then remembered utility should be insensitive to variations in the duration of the temporally extended experience. This prediction has been confirmed in both experimental and observational studies carried out with human subjects (Fredrickson & Kahneman, 1993; Kahneman et al., 1993; Redelmeier & Kahneman, 1996). For example, in retrospective evaluations of colonoscopy procedures that varied in duration from 4 - 67 minutes, aversiveness was not correlated with duration, but was strongly correlated with real-time ratings of both peak pain and pain at the end of the procedure (Redelmeier & Kahneman, 1996). Such insensitivity to duration has been called “duration neglect” (Fredrickson & Kahneman, 1993).

Figure 2 about here

Duration neglect has also been observed in the case of BSR (Gallistel, 1978; Mark & Gallistel, 1993; Shizgal & Matthews, 1977). Based on available data, Figure 2 depicts the simulated growth of instantaneous utility as a stimulation train is prolonged. The x-axis of the three-dimensional graph represents the aggregate firing rate produced by the stimulation in the neurons responsible for the rewarding effect; the higher the current or the frequency, the higher the aggregate firing rate. At each firing rate, the level of instantaneous utility climbs as the duration of the train is increased, eventually approaching asymptote. This saturation occurs quickly at high firing rates and more slowly at low ones (Mason & Milner, 1986). Duration neglect would be manifested by indifferent choice between two trains of the same strength (trains that produce the same aggregate firing rate) but different durations. This would be the case for values lying on the “plateau” of the depicted surface. Indeed, if the output of one and the same integrator were responsible both for the instantaneous and remembered utility of the stimulation, then the surface in figure 2 would not only describe the contribution of aggregate firing rate and duration to measures of choice, but also to measures of the resistance to the interruption of ongoing stimulation.

The translation of the instantaneous value of the stimulation-induced signal into a remembered utility has been modeled previously as the recording of the peak value (measuring the height of the plateau in Figure 2) (Gallistel, 1978; Gallistel et al., 1981). However, Kahneman's peak-and-end model makes the same prediction as a peak model in response to a steady, prolonged input (because the peak height is the same as the height at the end). Thus, further work is required to see which model works best in the case of BSR. Indeed, whereas the instantaneous utility at the end of an experience may be particularly important in assessing aversive states that one wishes to terminate, the value at the beginning may have a large bearing on the assessment of states that one wishes to initiate. Regardless of the relative contributions of beginnings and ends, the available data do suggest that the decision utility of BSR is computed in the spirit of the proposal by Kahneman et al. Rather than computing the temporal integral of instantaneous utility, the rat seems to apply a simple rule to a single exemplar value or a limited set thereof, thus showing profound neglect of duration. Determining the values that serve as the exemplars of instantaneous utility during different states and the rules used to combine these exemplars are important goals for future research.

In Figure 2, the signal responsible for BSR is portrayed as a unidimensional quantity that fluctuates in intensity over time as a function of aggregate firing rate and duration. In the following sections, the notion that a unidimensional signal is responsible for BSR is developed and the relationship of this signal to gustatory reward and the evaluative system are discussed.

Relationship between the utility of BSR and gustatory stimuli

A currency function expresses the value of different inputs on a common scale. Animals behave as if they routinely compute currency functions because they make orderly choices between complex, mutually exclusive alternatives, such as returning to the shelter of a nest or visiting a habitual foraging site. Each of these alternatives has multiple attributes germane to physiological regulation and risk. For example, these two options differ in the probability of finding food and water, losing body heat, and encountering a predator. Choosing the more

valuable option requires that the multidimensional representations be “boiled down” to a single common dimension (McFarland & Sibley, 1975), a common scale of utility.

In the experiments to be reviewed, the choices made by rats offered various reinforcers were recorded. By definition, these choices reflect the decision utilities of the available outcomes. In a later section, I will discuss the translation of remembered utility into decision utility. For now, let us assume that remembered utility was the only determinant of decision utility that varied in the experiments to be reviewed and hence, we can “see through” the translation process. Given this assumption and the portrayal provided above of how remembered utility is computed from instantaneous utility, the observed choices of the rats can be seen to reflect underlying changes in instantaneous utility.

To determine whether rats use a common currency to evaluate rewarding LH stimulation and a sucrose solution, Conover and I performed two types of experiments. First, we placed the rewarding LH stimulation in competition with the sucrose by presenting our subjects with a forced choice between them (Conover & Shizgal, 1994a). The strength of the BSR was varied across trials. Not surprisingly, the rats chose the sucrose in preference to the BSR when the strength of the electrical stimulation was below the threshold required to support responding in the absence of the sucrose. When the strength of the LH stimulation was set somewhat above this threshold, the rats continued to prefer the sucrose. In other words, the presence of the sucrose caused the rats to forgo trains of BSR for which they had worked rather vigorously in the absence of the gustatory stimulus. However, once the electrical stimulation was sufficiently strong, BSR was chosen exclusively in preference to the sucrose. Thus, the rats behaved as if they had selected the larger of two payoffs evaluated on a common scale.

In a subsequent experiment, we offered the rats a choice between BSR alone and a compound reward consisting of an intraoral infusion of sucrose and an equally preferred train of BSR. Five of the six rats preferred the compound reward to its electrical component alone. Thus, the effects of LH stimulation and sucrose summate in the computation of utility. Summation is possible only when the inputs share a common property that is registered by the system of

measurement. Given the above arguments and assumptions, the common property registered by our system of measurement, behavioral choice, is the ability to drive instantaneous utility to positive values.

The competition and summation experiments demonstrate that the LH stimulation and the sucrose have something important in common, much as Hoebel and others had proposed (Hoebel, 1969). However, two subsequent experiments demonstrate important differences between the gustatory and electrical rewards.

In one experiment (Conover, Woodside, & Shizgal, 1994), we increased the utility of a gustatory stimulus, a sodium chloride solution, by depleting the subjects of sodium. In the second experiment (Conover & Shizgal, 1994b), we decreased the utility of another gustatory stimulus, a sucrose solution, by allowing large quantities of this solution to accumulate in the gut. We reasoned that if the LH stimulation recreates the experience normally produced by a rewarding tastant, then manipulations that alter the utility of a tastant should have a similar effect on the BSR. The findings did not support such a hypothesis. Depleting the rats of sodium by administering a diuretic dramatically increased the utility of the saline solution without producing any observable change in the utility of BSR. Allowing large quantities of a sucrose solution to accumulate in the gut dramatically reduced the utility of this solution, in some cases rendering it aversive. However the same manipulation either failed to alter BSR or produced a much smaller reduction in the utility of the electrical reward than in the utility of the gustatory reward.

Figure 3 about here

Our results suggest that although a common signal represents the instantaneous utilities of the gustatory and electrical rewards, the LH electrode accesses the neural circuitry that computes this signal downstream from the point where gustatory stimuli are weighted by physiological feedback. Two models of how this could be arranged are shown in Figure 3. Routing physiological feedback to act at the inputs to the circuitry that computes the currency, enables behavior to contribute to the specificity of regulation. This can be seen by considering the alternative, a system where a currency function returns the relative utilities of sucrose and saline

on a common scale, and physiological feedback operates uniquely on the output values. In such a system, changes in sodium balance would alter the utility of both saline and sucrose solutions as would feedback from the gut following accumulation of a sucrose load. In contrast, if physiological feedback weights the inputs to the currency function, then the relative utilities of saline and sucrose solutions can be adjusted independently, thus biasing consumption in response to physiological needs.

Unidimensional versus multidimensional coding

In typical BSR experiments, neurons are excited within a relatively large region surrounding the electrode tip (Yeomans, 1990). The argument advanced by Shizgal and Conover (Shizgal & Conover, 1996) to tie the rewarding effect of BSR to the output of a currency function addresses the question of how electrical stimulation of a large population of cells could mimic a naturally occurring signal. Multiple coding dimensions are required to capture information about stimulus quality. If only a single dimension were available, then changes in quality would be indistinguishable from changes in intensity. This is why we see the world monochromatically under dim illumination, when only a single class of photoreceptors is activated. To represent multiple dimensions of information, some form of spatiotemporal coding is required. For example, the cells activated by the stimulus might be divided into multiple sub-populations, each sensitive to a particular quality (“labeled-line coding”), or a unitary population might produce different temporal patterns of activity in response to different stimulus qualities. In either extreme case or in mixtures thereof, it is unlikely that gross electrical stimulation would mimic the multidimensional code. Neurons that do not normally fire in concert would be activated simultaneously, and all the directly stimulated cells would fire with the same, rigid, stimulation-induced periodicity.

In contrast, the electrical stimulation could mimic the effect of a naturally-occurring stimulus if activity in the stimulated system represented only a single dimension of information. In such a system, an aggregate rate code suffices. In such a code, it matters neither which neuron fires nor when but only how many firings are produced by the entire population. The spatially

contiguous and temporally synchronous firing evoked by the electrode could produce the same number of firings in a system using an aggregate rate code as the spatially discontinuous and temporally asynchronous firing that is likely evoked by a natural stimulus. Thus the stimulation-induced activity would mimic the effect of the natural stimulus. Indeed, in studies of motion perception in unanesthetized monkeys, micro-stimulation of a population of neurons that appear to use aggregate firing rate to code a single perceptual dimension, the direction of visual motion, can mimic the effect of adding correlated motion to the elements of a visual stimulus (Newsome & Salzman, 1993).

As discussed above (see “counter model”), there is strong evidence that the decision utility of BSR and, by inference, its remembered and instantaneous utility, are derived from aggregate firing rate. An aggregate code is well-suited to represent values in a common currency, since, by definition, these values are arrayed along a single dimension. Thus, values derived from an aggregate code could be used to compare and combine the contributions to instantaneous utility of a draught of sucrose, a draught of saline, or a train of BSR.

The evaluative, perceptual, and timing channels

Inevitably, information is lost in boiling down a multidimensional representation of a stimulus to obtain a currency value. For example, one cannot recover the temperature, sweetness, or texture of a gustatory stimulus from a currency value representing its instantaneous utility. However, the information lost due to the collapsing of multiple dimensions is essential for identifying the stimulus and distinguishing it from others. Thus the circuitry that computes instantaneous utility must diverge from the perceptual circuitry subserving identification and discrimination. This divergence makes it possible to distinguish between the many different objects and outcomes that may happen to share the same utility. Similarly, in order to predict the time when the reinforcer will next be available, it is important to segregate information about when a reinforcer was encountered from information in the perceptual and evaluative channels. Thus, as depicted in Figure 4, information about reinforcers must be processed in at least three different ways.

Figure 4 about here

In this view, the perceptual channel tells the animal what and where the stimulus is, the evaluative channel returns the instantaneous utility of the stimulus, and the timer predicts when the reinforcer will next be available. Gross electrical stimulation of the evaluative channel could produce a meaningful signal if, as I have argued, information is encoded in the stimulated stage by the aggregate rate of firing. In contrast, gross electrical stimulation is unlikely to produce a meaningful, multidimensional signal in the perceptual channel because of the nature of the coding required. What about the response of the timer? It stands to reason that transitions in the state of many different channels would be accessible to the timer as “events.” If so, an abrupt change in the activity of the evaluative channel, such as the stimulation-induced perturbation responsible for BSR, may provide a sufficient input to support measurement of temporal intervals.

The perceptual channel is constructed to return facts about the world. Thus, it is equipped with constancies and normalization procedures that minimize the impact of changes in external or internal state on identification and discrimination. Of course, these constancies and normalization procedures are imperfect, and bandwidth limitations make it impossible for perception to be veridical. For example, subjective response varies non-linearly with changes in the strength of sensory stimuli, and it is possible to trick the perceptual system into producing illusions. Nonetheless, the system does a remarkably good job at estimating objective physical properties such as size, shape, distance, and reflectance.

The interval timer also appears to be designed to capture data about objective events. In scalar expectancy theory (Gibbon, 1977), the subjective measure of a temporal interval is a noisy scalar transform of the objective interval. Although the interval timer is less accurate than the circadian oscillator, it is highly flexible, operating over a huge temporal range and accommodating concurrent timing of multiple intervals with arbitrary stop and start times (Gibbon, Malapani, Dale, & Gallistel, 1997).

In contrast to the perceptual and timing channels, the evaluative channel operates without even a pretense of objectivity. External objects do not have an inherent worth independent of

present physiological and ecological conditions. For example, both the sign and magnitude of the instantaneous utility signal produced by a given stimulus can change as a function of physiological state (Cabanac, 1971); thus, a cool stimulus applied to the skin can be refreshing when one is overheated and unpleasant when one is hypothermic. The evaluative channel is constructed, not to return objective properties of stimuli, but rather to return a subjective estimate of the current significance of these properties.

In contemporary accounts of sensory information processing, the perceptual “channel” is often treated as a community of neural modules, each specialized to extract information of a particular kind such as color, form, movement, depth and texture. The evaluative channel can also be regarded as a specialized neural module, charged with the task of deriving another, more subjective, kind of information from the flow of sensation. It is not clear whether this module is also composed of a set of specialized processors, perhaps each linked to a given modality or combination thereof. If so, the outputs of all the evaluative processors that influence an ongoing course of action would have to be combined in order to compute instantaneous utility. (The continue/stop signal has only one degree of freedom.) Thus, the output of the evaluative system is treated here as unitary.

The circuitry responsible for interval timing would appear to constitute yet another module. Formal models of this stopwatch-like device have been developed and tested extensively in behavioral studies. Components of one such model, based on the scalar expectancy theory of timing (Church, 1984; Gibbon, 1977; 1995), have been linked to the activity of pharmacologically and anatomically characterized neural populations (Gibbon et al., 1997; Meck, 1996).

Natural reinforcers are processed by the perceptual, evaluative, and timing channels. In the following sections, the notion of remembered utility is generalized to reflect this parallel processing of information about reinforcers, and the roles of each of these channels in computing different dimensions of payoff are discussed.

Subjective dimensions of payoff

According to the view developed here, changes in the strength of the stimulation delivered in BSR experiments (e.g., changes in frequency or current) alter the aggregate firing rate of neurons that give rise to instantaneous utility. As shown in Figure 5, a stored record of this response to the change in stimulus strength is derived by applying a heuristic, such as the peak-and-end rule, to exemplar values of instantaneous utility, such as the beginning, peak and end. Variables controlling the strength of natural reinforcers, such as the concentration of a sucrose solution or the temperature of an air current, are viewed as acting analogously, with the exception that the impact of these variable is weighted by physiological state. Kahneman and his coworkers (Kahneman et al., 1997; Schreiber & Kahneman, submitted) have used the term “remembered utility” to refer to the stored record derived from exemplar values of instantaneous utility. In the remaining discussion, I will substitute the term “subjective intensity of the payoff” for remembered utility in labeling the stored appraisal of the variables that contribute to stimulus strength.

Figure 5 about here

Why introduce yet another term? I do this because decision utilities reflect not only the output of the evaluative channel subserving instantaneous utility but also the outputs of the timing and perceptual channels. Thus, the effects of reinforcers on future choices depend not only on their strength, but also on their rate, delay, amount, and kind. The notion of remembered utility proposed by Kahneman, Wakker, and Sarin (1997) and developed by Schreiber and Kahneman (submitted) is tied to the intensity dimension alone. A more general means of describing recorded payoffs is required if we are to capture the multidimensional contribution of reinforcers to decision utility.

To illustrate the need for a multidimensional treatment of reinforcement, consider the interaction of the rate and strength of reinforcement. Via preference tests, it can be shown that rats prefer highly concentrated solutions to less concentrated ones (Young, 1967). This difference in the intensity of the payoff can be offset by a compensatory change in its rate (Heyman & Monaghan, 1994): the allocation of time or responding to the two solutions can be equated by

making the less concentrated solution available more frequently than the highly concentrated one. The same relationship between stimulus strength and rate of reinforcement in determining behavioral allocation seems to hold when BSR, rather than a natural goal object, serves as the reinforcer (Hamilton, Stellar, & Hart, 1985). Similarly, weaker trains of stimulation are preferred equally to stronger trains when the rate at which the weaker trains are available is sufficiently high (Gallistel, 1991).

That both the rate and strength of reinforcers contribute to payoff suggests that the stored record of the reinforcer is multidimensional. The perceptual, timing, and evaluative channels not only process information about reinforcers in parallel, they also record their outputs in parallel. Payoff is then computed by combining the contents of the multidimensional record. In this view, illustrated in Figure 6, subjective intensity is the dimension of the stored record derived from the output of the evaluative channel. An output of the timing channel constitutes the second dimension. The nature of this stored quantity is a matter of debate. In one well-formulated proposal, this temporal dimension of the stored record contains a noisy measure of the inter-reinforcer interval (the inverse of reinforcement rate) (Gibbon, 1995). For convenience of phrasing I will use the term “subjective rate of payoff” to refer to this dimension of the stored record, leaving open the possibility that the stored quantity is not a rate *per se* but rather a measure from which a rate could be derived, such as an inter-reinforcement interval.

Figure 6 about here

In addition to subjective rate, the timer provides another quantity that contributes to subjective payoff: the delay between the reinforced response and the delivery of the reinforcer. Payoff appears to decline hyperbolically as the presentation of the reinforcer is delayed (Commons, Mazur, Nevin, & Rachlin, 1987; Mazur, 1986; Myerson & Green, 1995). This relationship appears to hold for BSR as well as for natural reinforcers (Mazur, Stellar, & Waraczynski, 1987).

The treatment of BSR presented here implies that a record consisting of a subjective intensity and the subjective weighting of rate and delay is sufficient for computing a payoff on

which a decision utility can be based. However, in the case of natural stimuli, additional dimensions contribute to payoff. For example, reinforcers, such as food pellets, may vary in mass. It stands to reason that the amount of a natural reinforcer would be recovered from perceptual information such as size and heft, and that unlike the case of intensity, the estimation of amount would be stable in the face of changes in physiological state. For example, one would hope that hunger would not alter one's judgments about the size of the fruit in a tree. Thus, an additional dimension of remembered payoff, the subjective weighting of amount, is likely returned by the perceptual channel. If BSR is not accompanied by a meaningful signal in the perceptual channel, then the information in this cell of the payoff record is likely to be absent or indecipherable.

The contribution of payoff to decision utility would appear to involve at least two stages of processing. First, a stored record is obtained by mapping physical dimensions such as strength, rate, delay, and amount into corresponding subjective ones. Decision utility would appear to reflect the result of performing a combinatorial operation on the quantities in the stored record of payoff. Accounts of matching, to be discussed in the next section, tend to treat the combinatorial operation in question as multiplication (Baum & Rachlin, 1969; Davison & McCarthy, 1988). Figure 7 depicts multiplicative combination of intensity and rate in computing the subjective payoff provided by a train of rewarding stimulation.

Figure 7 about here

Matching: translation of subjective payoff into decision utility

Imagine a pair of exquisite but idiosyncratically managed restaurants. The quality of the cuisine may be truly outstanding, and the philanthropic proprietors demand no remuneration other than the commitment of time by the diners. Each restaurant is open a certain number of times per month on the average, a rate that may or may not differ from the accessibility of the competing establishment. Although the average rate at which each restaurant opens is constant over time, the interval between openings varies randomly. Openings are unannounced, thus keeping the clientele guessing as to the date and time when they can gain entry. Sometimes, a client arrives to find the chosen restaurant already open; on other occasions, the restaurant is closed at the time of arrival,

and the client may either wait until the restaurant opens or leave beforehand. Once a seat at a table has been secured, the diners face a delay until the serving of the first course that may differ in the two establishments,. Finally, both chefs are experimenting with different portion sizes; currently, one leans towards “cuisine minceure” and the other towards the fashion of a Chicago steak house.

The matching law was formulated to describe the allocation of behavior in experimental settings roughly analogous to the competing restaurants. Lest the reader find the capricious scheduling too bizarre to take seriously, I should point out that the unpredictable availability of a reinforcer might well seem more realistic to people living in the manner of our ancestors. The traditional Inuit hunter did not expect a seal to visit a particular breathing hole in the ice at any designated time, yet he derived an estimate of the average frequency of visits and allocated his time accordingly.

In the terminology of operant conditioning, the diners in the above example are presented with concurrent variable-interval schedules of reinforcement According to the strict form of the matching law (Davison & McCarthy, 1988; de Villiers, 1977; Herrnstein, 1961; Herrnstein, 1970; Williams, 1988), they will allocate their time and visits in proportion to the relative payoffs provided by the two restaurants. These payoffs are calculated by multiplicative combination of the subjective intensity (the “goodness” of the food) with the subjective weightings of the rate of opening, delay of meal onset and portion size. In the terms employed here, the matching law translates the multidimensional records of subjective payoffs into decision utilities. Under the strict form of the matching law, relative decision utility is proportional to relative subjective payoff.

Figure 8 about here

The subject performing on concurrent variable-interval schedules can be portrayed as repeatedly flipping a biased coin, with the bias reflecting the relative payoffs provided by the two reinforcers (Gibbon, 1995; Heyman, 1988; Heyman & Goodman, submitted). If so, the relative payoffs will be reflected in the relative allocations of time to the two schedules. Given strict

matching and multiplicative combination of subjective intensity and rate, the ratio of the subjective intensities of two reinforcers can then be calculated from the observed ratios of reinforcement rates and time allocation. This logic was used by Miller (1976) to measure, in pigeons, the relative intensity of the payoffs provided by three different types of seeds and by Gallistel and his students (Gallistel, 1991; Leon & Gallistel, 1992; Simmons & Gallistel, 1994) to measure how the subjective intensity of BSR grows as a function of the strength of electrical stimulation in rats. Gallistel's group found that as the stimulation strength rises above threshold, the subjective intensity of the payoff climbs steeply, initially approximating a power function; the growth eventually slows and levels off as stimulation strength is increased to ever higher levels. In Figures 2, 5, 7 and 8, the growth of BSR as a function of the aggregate rate of stimulation-induced firing is modeled as a logistic, thus capturing the steep initial rise, the later deceleration, and the eventual leveling off.

Knowing the form and parameters of the intensity-growth function could serve as a powerful constraint in interpreting recordings of neural activity. For example, if one wished to argue that a particular population of neurons encodes the intensity of the payoff produced by the rewarding stimulation, one would have to demonstrate that some attribute of activity in this population corresponds to the form and parameters of the intensity-growth function.

Beyond strict matching: contribution of economic constraints to decision utility

The above discussion of matching was confined to cases in which the two competing reinforcers are of the same kind (e.g., food) and where the reinforcers are available outside the test environment. When reinforcers of different kinds are pitted against each other in choice experiments or when a natural reinforcer is available uniquely in the test environment, the strict form of the matching law may no longer account gracefully for the translation of subjective payoffs into decision utility. Two additional pieces of information appear to contribute to the computation of decision utility: the category to which the reinforcer belongs (its kind) and the environmental distribution of the reinforcer. The role of the perceptual channel in providing this information is highlighted by experiments in which BSR competes with natural reinforcers.

The needs addressed by behavioral means are many and varied. Thus, humans and other animals seek “goods” of different kinds. If a sufficiently broad time frame is adopted, choices between alternatives are rarely constrained to single decisions between mutually exclusive options. Under such circumstances, the problem of adaptive choice is to select the best “bundle” of goods rather than the single item with the highest value in a common currency. Research in behavioral economics suggests that when available goods are not of the same kind, calculating the utility of the bundle will usually require an operation more complex than simply summing the utilities of the individual items (Kagel, Battalio, & Green, 1995; Rachlin, Green, Kagel, & Battalio, 1976; Rachlin, Kagel, & Battalio, 1980).

The relationships between the items in the bundle can be arrayed on a continuum. At one extreme are goods that are entirely substitutable, like two brands of cola to an indifferent consumer. At the other extreme are goods that are complements, such as bicycle frames and bicycle wheels or left shoes and right shoes. In the case of perfect substitutes, the utility of the bundle is simply the sum of the utilities of the constituent goods. In contrast, the more complementary the constituents, the less each is worth individually in comparison with the utility of the bundle.

The distinction between substitutes and complements has profound behavioral implications. For example, if the relative prices of two highly substitutable goods are changed and the budget of the consumer is adjusted so as to make possible the purchase of the same quantities as had been acquired at the former prices, consumption shifts toward the cheaper good. In contrast, complementary goods tend to be consumed in a fixed ratio, which is insensitive to changes in relative price.

When substitutability is taken into account, the simple form of the matching law no longer suffices. For example, responding for the reinforcer on the richer of two schedules will increase when the schedule for a perfect substitute is made leaner (increasing its “price”), as the matching law predicts. However, responding on the richer schedule will decrease when the schedule for a perfect complement is made leaner. Rachlin, Kagel and Battalio (Rachlin et al., 1980) have shown

how a generalization of the simple form of the matching law can be interpreted to incorporate substitutability. Each of the ratios reflecting the subjective weighting of rate, delay, amount, and strength is first raised to an exponent before scalar combination. For perfect substitutes, the exponent is one, and the equation reduces to the simple form of the matching law. According to this modification to the matching law, the translation of subjective payoff into decision utility depends on the substitutability of the reinforcers.

Substitutability of BSR and natural reinforcers. Results of an experiment carried out by Green and Rachlin to measure the substitutability of rewarding LH stimulation, food, and water (Green & Rachlin, 1991) can be interpreted in terms of the role of the perceptual channel in computing decision utility. As expected, they found that food and water were poor substitutes. Nonetheless, BSR was highly substitutable for both food and water.

Presumably, dry food and water are poor substitutes because they fulfill non-overlapping physiological needs. If so, does the high substitutability of BSR with both food and water imply that the signal injected by the electrode mimics signals specific both to energy balance and fluid balance? If this were the case, one would expect that BSR would be less substitutable for food and water than for itself. This isn't what Green and Rachlin found. BSR triggered by one lever was about as substitutable for BSR triggered by the other lever as for food or water.

One interpretation suggested by Green and Rachlin is that BSR acts as a "general" reinforcer. This is reminiscent of the coding argument advanced by Shizgal and Conover (Shizgal & Conover, 1996). We proposed that any activity produced by the rewarding stimulation in the perceptual system would tend to be "noisy" and would be unlikely to mimic the signature of a naturally occurring goal object. Moreover, our results suggest that the rewarding stimulation acts downstream from the point where physiological feedback weights gustatory rewards. Thus, there may well be no way for the rat to determine the regulatory system or the class of goal object to which the stimulation is germane. However, our data suggest that the evaluative channel produces a unidimensional currency signal in response to rewarding brain stimulation, a sucrose solution and a saline solution. If this were also the case for water and for the food employed by Green and

Rachlin, then these goods would also substitute well for BSR. In contrast, food and water would register distinct signals in the perceptual channel, making it possible to incorporate the identity, and hence, the low substitutability, of these two reinforcers when computing the relative decision utilities of the activities that produce these two goods.

Elasticity of demand for BSR and natural reinforcers. Most experiments that employ food reinforcement are conducted under the conditions of an “open economy:” food is made available during the relatively short test sessions and is also available during at least some of the much longer period the subject spends in its home cage. Often, the body weight of the subjects is maintained at some proportion of the free-feeding value. Thus, the supplementary food given to the subjects in the home environment can bring total intake to a fixed level when added to what was earned in the test environment. In such circumstances, “demand” (the number of reinforcers earned in the test environment) is said to be highly “elastic,” changing steeply in response to variations in “price” (the number of responses required to obtain a single reinforcement). For example, the higher the cost of food in the experimental situation, the less food the animal earns there (Hursh, 1980). In contrast, when the economy is “closed,” and the reinforcer is available uniquely in the test environment, demand for food becomes highly *inelastic* over an appreciable range of prices (Collier, Johnson, Hill, & Kaufman, 1986; Hursh, 1980). When the subject must fend for itself, without the benefit of supplements, response rates tend to increase with price so that consumption is largely defended.

Although it is tempting to attribute the inelastic demand observed in a closed economy to an accumulating effect of deprivation, the data do not fit this hypothesis gracefully. Over the range of prices within which demand remains highly inelastic in such experiments, subjects can maintain their body weight or even increase it (Collier, 1983). If so, some aspect of the economic circumstances appears to be altering the translation of subjective payoff into decision utility.

According to the single-operant version of the matching law (Herrnstein, 1970), subjects presented with a single experimenter-supplied reinforcer choose between working for it and performing alternative activities, such as the opportunity to groom, rest, or explore. Herrnstein’s

formulation predicts that increasing the price of the experimenter-supplied reinforcer will shift the allocation of behavior away from that reinforcer and towards “everything else.” The effect of price on performance for food in a closed economy violates this form of the matching law: contrary to prediction, relative allocation of behavior to natural reinforcers is no longer proportional to relative payoff.

The situation is different in the case of BSR. In the great majority of BSR experiments, the economy is closed; no stimulation is available in the home cage. Yet demand for BSR is highly elastic; the larger the number of responses required to earn a reinforcement or the more stringent the limit on maximum consumption per unit time, the lower the allocation of behavior to responding for BSR (Druhan, Levy, & Shizgal, 1993; Fouriez, Emdin, & Beaudoin, 1996; Hamilton et al., 1985). Thus, as Hursh and Natelson (1981) have demonstrated, closing the economy produces very different changes in performance for BSR and food: demand for BSR remains highly elastic whereas demand for food is highly inelastic.

The effect of closing the economy on responding for food and water reinforcement has been interpreted to suggest that the subject learns the relative availability of the reinforcer in the different environments to which it is exposed (Collier, 1983), much as optimal foraging theorists have postulated (Charnov, 1976; Stephens & Krebs, 1986). Such a representation would appear to store information from the perceptual channel in a spatial or spatiotemporal context (e.g., food is available in the test chamber, but not in the home cage) and to contribute, along with the subjective weightings of rate, delay, amount, and strength, to determining decision utility. The situation for BSR seems far simpler: demand for BSR appears unaffected by whether or not the economy is open (although direct experimental confirmation of this prediction is lacking). Perhaps this is so because the “noisy” signal produced by rewarding brain stimulation in the perceptual channel cannot be attributed to a particular natural reinforcer.

Expectancy

In the fanciful example described above, the prospective diners could not predict the time when either of the restaurants would next open. Although the availability of resources in the

natural world may also be unpredictable, there are circumstances in which events occur in reliable sequences. For example, after a flower has been drained by a hummingbird, the nectar will tend to accumulate at a characteristic rate, and the bird will adjust the timing of its return accordingly (Gallistel, 1990). Subjects working on fixed-interval schedules of reinforcement in laboratory experiments manifest a pause in responding after delivery of the reinforcer; responding then resumes as the time to the next scheduled delivery of reinforcement approaches (Gibbon, 1977). In such circumstances, decision utility is adjusted dynamically to reflect information provided by the interval timer concerning the scheduling of reinforcement. The probability that the animal will direct its behavior towards the goal shifts from low to high as the predicted time of reinforcement approaches. This adjustment in behavior is anticipatory and thus, the animal can be said to have formed an expectation of future payoff.

In scalar expectancy theory, the average level of expectancy is increased by a large recent payoff (Gibbon, 1977). It stands to reason that this increase might be specific to the expectancy of future payoffs of the same kind as the large, recent one. This proposal is in the spirit of Loewenstein's (1996) treatment of the focusing effects produced by sensations and states that have large instantaneous utilities. Such sensations and states are portrayed as "crowding out" consideration of alternate reinforcers.

With this view of expectancy in mind, let us return to the vignette at the beginning of this chapter. Towards the end of the rat's confinement in the start box, it approaches the barrier blocking access to the alley. Such behavior is much more pronounced on trials when stimulation is delivered in the start box than on trials when it is not; on the trials preceded by stimulation, the rat not only approaches the barrier but makes frenzied attempts to climb over it. This potentiation of anticipatory acts cannot be due simply to stimulation-induced arousal because the same start-box stimulation fails to galvanize behavior when the rat has learned that additional stimulation is no longer available in the goal box. Rather, the experience of a large, recent payoff seems to rescale expectancy (Sax & Gallistel, 1990), boosting its magnitude at each point in time as the delay to accessing the alley elapses. However, when the rat has learned that reinforcement is no longer

available in the goal box, there is, in effect, no expectancy for the start-box stimulation to rescale, and the rat's behavior in the start box becomes nonchalant.

The effect of a large, recent payoff on decision utility is illustrated by two experiments in which thirsty rats chose between BSR and water (Deutsch, Adams, & Metzner, 1964; Wasserman, Gomita, & Gallistel, 1982). Shortly following delivery of pretrial stimulation, the rats chose BSR in preference to water. When the pretrial stimulation was omitted, the preference reversed, and the rats opted for the water instead of the BSR. The pretrial stimulation was very similar to the stimulation offered as a reward. In contrast, the stimulation is unlikely to have mimicked the perceptual experience produced by the water (see "Unidimensional versus multidimensional coding"). According to the argument sketched out above, the expectancy boosted by the large pretrial payoff will be directed at the reward with the neural signature most similar to that of the pretrial stimulation. This augmented expectancy would boost the decision utility of BSR, thus biasing choice. It would be interesting to determine whether pretrial exposure to water would produce a complementary effect. More generally, the dependence of decision utility on the interaction between the experience, expectation, and recording of payoffs deserves additional attention. Rewarding brain stimulation may prove to be a valuable tool for investigating such interactions.

Coda

Vigorous efforts are underway to identify the components of the neural circuitry responsible for BSR. Although the quarry has long proved elusive, new methods for visualizing candidate neurons (Arvanitogiannis, Flores, Pfaus, & Shizgal, 1996a; Arvanitogiannis, Flores, & Shizgal, 1997; Flores, Arvanitogiannis, & Shizgal, 1997) and for measuring behavioral effects of drugs and lesions (Arvanitogiannis, Waraczynski, & Shizgal, 1996b; Shizgal, Conover, & Arvanitogiannis, 1996) offer hope for better hunting. Once these cells are found, we will have new questions to ask them as a result of this attempt to align conceptions of utility that have grown out of the study of BSR and natural reinforcers in laboratory animals with ideas derived from the study of evaluation and choice in humans. It will be particularly interesting to test the

hypothesis that the neurons responsible for the signal recorded as the subjective intensity of the payoff also give rise to the instantaneous utility of the stimulation. By recording from these neurons in awake, behaving subjects, it should prove possible to test many of the hypotheses discussed here concerning the role of these cells in the computation of utility. Among these hypotheses is the proposal that the output of the directly activated neurons underlying BSR can gain access to working memory under attentional control. Investigating this hypothesis could shed light on how signals fundamental to the experience of pleasure gain access to awareness.

Figure Captions

Figure 1: The counter model. Action potentials elicited in the directly activated neurons responsible for BSR impinge on a neural circuit that integrates their effects over time and space. The output of this integrator is determined by the aggregate rate of firing during a fixed time window. Thus, firing two neurons four times each produces the same output as firing four neurons twice each. (In addition to triggering action potentials that propagate to the synaptic terminals, the stimulation also triggers action potentials that propagate “backwards,” toward the cell body. These “antidromic” action potentials, shown in gray, will have no behavioral effect unless they invade another axonal branch.)

Figure 2: Growth of instantaneous utility as a function of stimulation strength and duration. The stronger the stimulation, the higher the aggregate rate of firing in the directly stimulated neurons responsible for the rewarding effect. Three relationships are depicted by the three-dimensional graph. With aggregate firing rate held constant, instantaneous utility climbs as the duration of the input is prolonged, eventually leveling off. This leveling off is responsible for the “duration neglect” that has been reported in BSR experiments (Gallistel, 1978; Mark & Gallistel, 1993; Shizgal & Matthews, 1977). With duration held constant, instantaneous utility climbs steeply as the aggregate firing rate is increased and then levels off. A logistic growth function has been used to simulate this effect. The third relationship is depicted in the projected contour map. The outlines of successive horizontal sections through the three-dimensional structure have been projected onto this plane. Each contour line gives the combinations of aggregate firing rate and train duration that raise instantaneous utility to a given “altitude.” The contour lines follow the hyperbolic form first described by Gallistel (1978). Changing the “altitude” at which the cross-section is taken shifts the curve along the axis representing the logarithm of the firing rate but does not change the curvature. Plotting the growth of instantaneous utility as a function of both aggregate firing rate and train duration illustrates an important consequence of the parallelism of the contour lines: the rate at which instantaneous utility grows with train duration increases as a function of aggregate firing rate. At high aggregate firing rates, instantaneous utility approaches

asymptote very quickly; at low firing rates, much more time is required for instantaneous utility to level off. Results consistent with this relationship have been reported by Mason and Milner (1986).

Figure 3: Alternative schemes for combining the rewarding effects of LH stimulation and gustatory stimuli. On the basis of experiments by Conover et al. (Conover & Shizgal, 1994b; Conover et al., 1994), signals that give rise to gustatory reward are weighted by physiological feedback prior to their combination with the signals that give rise to BSR. In the upper right panel, the two rewards are combined by passing the gustatory reward through the population of neurons from which the stimulating electrode samples. Thus, the post-synaptic effects of the gustatory and electrical rewards are integrated by a common circuit. In the lower panel, the gustatory and electrical reward signals are integrated separately before they are combined and relayed to the choice mechanism. (Modified from (Shizgal & Conover, 1996).)

Figure 4: Parallel channels process information about goal objects in real time. The perceptual channel returns the identity, location, and amount of the goal object, whereas a stopwatch-like channel marks the time when the goal object was encountered. The evaluative channel steers ongoing behavior so as to maintain or terminate contact with the goal object. Given sufficient allocation of attention and working memory, the output of the evaluative channel may be manifested in hedonic experience as pleasure or suffering.

Figure 5: “Representation by exemplar” (Schreiber & Kahneman, submitted) in computing the subjective intensity of BSR. The quantity recorded in memory (the subjective intensity) is derived from exemplar values of instantaneous utility, such as the peak (Gallistel, 1978; Gallistel et al., 1981) or the peak and end (Kahneman et al., 1993; Redelmeier & Kahneman, 1996). These exemplar values are independent of the temporal integral of instantaneous utility. Thus, subjects working for BSR manifest duration neglect: once instantaneous utility has reached the plateau of the plotted surface, further increases in duration fail to increase the remembered subjective intensity.

Figure 6: Recording the output of the parallel information-processing channels. Stored information from all three channels contributes to payoff. Information derived from the perceptual channel indicates “kind” (is the goal object a source of food, water, or salt?) as well as amount. Estimates of the encounter rate and the delay between a successful response and delivery of a reinforcer are derived from the output of the stopwatch timer. The evaluative channel contributes an estimate of subjective intensity (see Fig. 4) to the payoff record.

Figure 7: Computation of the subjective payoff provided by a train of rewarding stimulation. The left-hand portion of the figure, reproduced from Figures 1 and 2, shows how the instantaneous utility of the rewarding stimulation is derived from the aggregate firing rate in the directly stimulated stage of the underlying neural circuit. Via the principle of representation by exemplar, instantaneous utility is transformed into the subjective intensity of the payoff, one of the dimensions of the stored record of subjective payoff. Two possible rules for carrying out this transformation are shown: peak-and-end averaging and peak detection. A second dimension of the stored record, the subjective rate of payoff, is provided by an interval timer. On the basis of research on operant matching, the combinatorial operation for combining these two dimensions is shown as multiplication.

Figure 8. Translation of subjective payoff into behavior according to Herrnstein’s treatment of operant performance for a single experimenter-controlled reinforcer. According to Herrnstein’s view, the payoff obtained by working for the experimenter-controlled reinforcer is compared to the payoff from competing activities such as grooming, exploring, and resting (“everything else”). The allocation of behavior to the experimenter-controlled reinforcer is determined by the payoff it provides as a proportion of the sum of all payoffs available in the test environment. This view of behavioral allocation runs into difficulty when the subject works for an essential natural reinforcer unavailable outside the test environment or when the subject chooses between two natural reinforcers of different kinds. However, neither of these restrictions apply in the case of BSR. Reasons why performance for BSR and for natural reinforcers differ in these respects are discussed in the text.

References

- Arvanitogiannis, A., Flores, C., Pfaus, J. G., & Shizgal, P. (1996a). Increased ipsilateral expression of Fos following lateral hypothalamic self-stimulation. *Brain Research*, **720**, 148-154.
- Arvanitogiannis, A., Flores, C., & Shizgal, P. (1997). Fos-like immunoreactivity in the caudal diencephalon and brainstem following lateral hypothalamic self-stimulation. *Behavioural Brain Research*, **88**(2), 275-279.
- Arvanitogiannis, A., Waraczynski, M., & Shizgal, P. (1996b). Effects of excitotoxic lesions of the basal forebrain on MFB self-stimulation. *Physiology & Behavior*, **59**(4/5), 795-806.
- Baars, B. J. (1988). *A Cognitive theory of consciousness*. Cambridge: Cambridge University Press.
- Baum, W. M., & Rachlin, H. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior*, **12**, 861-874.
- Bentham, J. (1996 (originally published: 1789)). *An introduction to the principles of morals and legislation*. Oxford: Clarendon Press.
- Bielajew, C., & Shizgal, P. (1980). Dissociation of the substrates for medial forebrain bundle self-stimulation and stimulation-escape using a two-electrode stimulation technique. *Physiology and Behavior*, **25**, 707-711.
- Bishop, M. P., Elder, S. T., & Heath, R. G. (1963). Intracranial self-stimulation in man. *Science*, **140**, 394-6.
- Bower, G. H., & Miller, N. E. (1958). Rewarding and punishing effects from stimulating the same place in the rat's brain. *Journal of Comparative and Physiological Psychology*, **51**, 669-674.
- Boyd, E. S., & Gardiner, L. C. (1962). Positive and negative reinforcement from intracranial stimulation of a teleost. *Science*, **136**, 648-9.
- Cabanac, M. (1971). Physiological role of pleasure. *Science*, **173**, 1103-1107.

- Cabanac, M. (1992). Pleasure: the common currency. *Journal of Theoretical Biology*, **155**, 173-200.
- Cabanac, M., & LeBlanc, J. (1983). Physiological conflict in humans: fatigue vs. cold discomfort. *American Journal of Physiology*, **244**, R621-R628.
- Charnov, E. L. (1976). Optimal foraging: the marginal value theorem. *Theoretical Population Biology*, **9**, 129-136.
- Church, R. M. (1984). Properties of the internal clock. In J. Gibbon & L. Allan (Eds.), *Timing and Time Perception* (Vol. 423, pp. 566-582). New York: New York Academy of Sciences.
- Collier, G. H. (1983). Life in a closed economy: The ecology of learning and motivation. In M. D. Zeller & P. Harzem (Eds.), *Advances in Analysis of Behaviour* (Vol. 3, pp. 223-274). New York: John Wiley & Sons.
- Collier, G. H., Johnson, D. F., Hill, W. L., & Kaufman, L. W. (1986). The economics of the law of effect. *Journal of the Experimental Analysis of Behavior*, **46**, 113-136.
- Commons, M. L., Mazur, J. E., Nevin, J. A., & Rachlin, H. (Eds.). (1987). *Quantitative Analysis of Behavior: The Effects of Delay*. (Vol. 5). Cambridge, MA: Ballinger.
- Conover, K. L., & Shizgal, P. (1994a). Competition and summation between rewarding effects of sucrose and lateral hypothalamic stimulation in the rat. *Behavioral Neuroscience*, **108**(3), 537-548.
- Conover, K. L., & Shizgal, P. (1994b). Differential effects of postingestive feedback on the reward value of sucrose and lateral hypothalamic stimulation in the rat. *Behavioral Neuroscience*, **108**(3), 559-572.
- Conover, K. L., Woodside, B., & Shizgal, P. (1994). Effects of sodium depletion on competition and summation between rewarding effects of salt and lateral hypothalamic stimulation in the rat. *Behavioral Neuroscience*, **108**(3), 549-558.
- Davison, M., & McCarthy, D. (1988). *The Matching Law*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- de Villiers, P. (1977). Choice in concurrent schedules and a quantitative formulation of the law of effect. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of Operant Behavior* (pp. 233-287). Englewood Cliffs, NJ: Prentice Hall.
- Deutsch, J. A., Adams, D. W., & Metzner, R. J. (1964). Choice of intracranial stimulation as a function of delay between stimulations and strength of competing drive. *Journal of Comparative and Physiological Psychology*, **57**, 241-243.
- Distel, H. (1978). Behavior and electrical brain stimulation in the green iguana, *Iguana iguana* L. II. Stimulation effects. *Experimental Brain Research*, **31**(3), 353-367.
- Druhan, J. P., Levy, M., & Shizgal, P. (1993). Effects of varying reinforcement schedule, reward current, and pretrial priming stimulation on discrete-trial performance for brain stimulation reward. *Psychobiology*, **21**(1), 37-42.
- Edmonds, D. E., & Gallistel, C. R. (1974). Parametric analysis of brain stimulation reward in the rat: III. Effect of performance variables on the reward summation function. *Journal of Comparative and Physiological Psychology*, **87**, 876-883.
- Flores, C., Arvanitogiannis, A., & Shizgal, P. (1997). Fos-like immunoreactivity in forebrain regions following self-stimulation of the lateral hypothalamus and the ventral tegmental area. *Behavioural Brain Research*, **87**(2), 239-251.
- Fouriezos, G., Emdin, K., & Beaudoin, L. (1996). Intermittent rewards raise self-stimulation thresholds. *Behavioural Brain Research*, **74**, 57-64.
- Frank, R. A., & Stutz, R. M. (1984). Self-deprivation: A review. *Psychological Bulletin*, **96**(2), 384-393.
- Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology*, **65**(1), 45-55.
- Gallistel, C. R. (1978). Self-stimulation in the rat: Quantitative characteristics of the reward pathway. *Journal of Comparative and Physiological Psychology*, **92**, 977-998.
- Gallistel, C. R. (1980). *The organization of action: a new synthesis*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, Massachusetts: MIT Press.
- Gallistel, C. R. (1991). Measuring the subjective magnitude of brain stimulation reward by titration with rate of reward. *Behavioral Neuroscience*, **105**(6), 913-925.
- Gallistel, C. R., Shizgal, P., & Yeomans, J. S. (1981). A portrait of the substrate for self-stimulation. *Psychological Review*, **88**, 228-273.
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, **84**(3), 279-325.
- Gibbon, J. (1995). Dynamics of time matching: arousal makes better seem worse. *Psychonomic Bulletin & Review*, **2**(2), 208-215.
- Gibbon, J., Church, R. M., Fairhurst, S., & Kacelnik, A. (1988). Scalar expectancy theory and choice between delayed rewards. *Psychological Review*, **95**(1), 102-114.
- Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. (1997). Towards a neurobiology of temporal cognition: advances and challenges. *Current Opinion in Neurobiology*, **7**(2), 170-184.
- Goldman-Rakic, P. S. (1996). Regional and cellular fractionation of working memory. *Proceedings of the National Academy of Sciences, USA*, **93**, 13473-13480.
- Green, L., & Rachlin, H. (1991). Economic substitutability of electrical brain stimulation, food, and water. *Journal of the Experimental Analysis of Behavior*, **55**, 133-143.
- Hamilton, A. L., Stellar, J. R., & Hart, E. B. (1985). Reward, performance, and the response strength method in self-stimulating rats: validation and neuroleptics. *Physiology & Behavior*, **35**, 897-904.
- Heath, R. G. (1964). Pleasure response of human subjects to direct stimulation of the brain: Physiologic and psychodynamic considerations. In R. G. Heath (Ed.), *The Role of Pleasure in Behavior* (pp. 219-243). New York: Harper and Row.
- Heller, J. (1961). *Catch-22*. New York: Simon & Schuster.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, **4**, 267-272.

- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, **13**(2), 243-266.
- Heyman, G. (1988). How drugs affect cells and reinforcement affects behavior: formal analogies. In M. L. Commons, R. M. Church, J. R. Stellar, & A. R. Wagner (Eds.), *Biological Determinants of Reinforcement* (Vol. VII, pp. 157-182). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Heyman, G. H., & Goodman, J. B. (submitted). Matching as an elementary behavioral principle: a Markov analysis of preference in concurrent choice procedures. .
- Heyman, G. M., & Monaghan, M. M. (1994). Reinforcer magnitude (sucrose concentration) and the matching law theory of response strength. *Journal of the Experimental Analysis of Behavior*, **61**, 505-516.
- Hoebel, B. G. (1969). Feeding and self-stimulation. *Annals of the New York Academy of Sciences*, **157**, 758-778.
- Hursh, S. R. (1980). Economic concepts for the analysis of behavior. *Journal of the Experimental Analysis of Behavior*, **34**, 219-238.
- Hursh, S. R., & Natelson, B. H. (1981). Electrical brain stimulation and food reinforcement dissociated by demand elasticity. *Physiology & Behavior*, **26**, 509-515.
- Johnson-Laird, P. N. (1988). *The computer and the mind*. Cambridge, Massachusetts: Harvard University Press.
- Kagel, J. K., Battalio, R. C., & Green, L. (1995). *Economic choice theory: an experimental model of animal behavior*. Cambridge: Cambridge University Press.
- Kahneman, D. (1994). New challenges to the rationality assumption. *Journal of Institutional and Theoretical Economics*, **150**(1), 18-36.
- Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: adding a better end. *Psychological Science*, **4**(6), 401-405.
- Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics*, **112**(2), 375-405.

- Lamb, R. J., Preston, K. L., Schindler, C. W., Meisch, R. A., Davis, F., Katz, J. L., Henningfield, J. E., & Goldberg, S. R. (1991). The reinforcing and subjective effects of morphine in post-addicts: a dose-response study. *Journal of Pharmacology and Experimental Therapeutics*, **259**(3), 1165-1173.
- LeDoux, J. (1996). *The emotional brain*. New York: Simon & Schuster.
- Leon, M., & Gallistel, C. R. (1992). The function relating the subjective magnitude of brain stimulation reward to stimulation strength varies with site of stimulation. *Behavioural Brain Research*, **52**, 183-193.
- Lepore, M., & Franklin, K. B. J. (1992). Modelling drug kinetics with brain stimulation: Dopamine antagonist increase self-stimulation. *Pharmacol Biochem Behav*, **41**, 489-496.
- Lilly, J. C., & Miller, A. M. (1962). Operant conditioning of the bottlenose dolphin with electrical stimulation of the brain. *Journal of Comparative and Physiological Psychology*, **55**, 73-79.
- Loewenstein, G. (1996). Out of control: visceral influences on behavior. *Organizational Behavior and Human Decision Processes*, **65**(3), 272-292.
- Macfarlane, D. B. (1954, March 12, 1954). McGill opens vast new research field with brain "pleasure area" discovery. *The Montreal Star*, pp. 1-2.
- Mark, T. A., & Gallistel, C. R. (1993). Subjective reward magnitude of medial forebrain stimulation as a function of train duration and pulse frequency. *Behavioral Neuroscience*, **107**(2), 389-401.
- Mason, P., & Milner, P. (1986). Temporal characteristics of electrical self-stimulation reward: fatigue rather than adaptation. *Physiology & Behavior*, **36**, 857-860.
- Mazur, J. E. (1986). Choice between single and multiple delayed reinforcers. *Journal of the Experimental Analysis of Behavior*, **46**, 67-78.
- Mazur, J. E., Stellar, J. R., & Waraczynski, M. (1987). Self-control choice with electrical stimulation of the brain. *Behavioural Processes*, **15**, 143-153.

- McFarland, D. J., & Sibley, R. M. (1975). The behavioural final common path. *Philosophical Transactions of the Royal Society of London B*, **270**, 265-293.
- Meck, W. H. (1996). Neuropharmacology of timing and time perception. *Cognitive Brain Research*, **3**, 227-242.
- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*: Holt, Rinehart and Winston, Inc.
- Miller, H. L. (1976). Matching-based hedonic scaling in the pigeon. *Journal of the Experimental Analysis of Behavior*, **26**, 335-347.
- Morgan, C. W., & Mogenson, G. J. (1966). Preference of water-deprived rats for stimulation of the lateral hypothalamus and water. *Psychonomic Science*, **6**, 337-338.
- Myerson, J., & Green, L. (1995). Discounting of delayed rewards: models of individual choice. *Journal of the Experimental Analysis of Behavior*, **64**, 263-276.
- Newsome, W. T., & Salzman, C. D. (1993). The neuronal basis of motion perception. *Ciba Foundation Symposium*, **174**, 217--246.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review*, **84**(3), 231-259.
- Olds, J. (1956). Pleasure centers in the brain. *Scientific American*, **195**, 105-116.
- Olds, J. (1958). Self-stimulation of the brain. *Science*, **127**, 315-324.
- Olds, J., & Milner, P. M. (1954). Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*, **47**, 419-427.
- Pfaffmann, C., Norgren, R., & Grill, H. J. (1977). Sensory affect and motivation. *Annals of the New York Academy of Sciences.*, **290**, 18-34.
- Porter, R. W., Conrad, D. G., & Brady, J. V. (1959). Some neural and behavioral correlates of electrical self-stimulation of the limbic system. *Journal of the Experimental Analysis of Behavior*, **2**, 43-55.

- Rachlin, H., Green, L., Kagel, J. H., & Battalio, R. C. (1976). Economic demand theory and psychological studies of choice. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10, pp. 129-154). New York: Academic Press.
- Rachlin, H., Kagel, J. H., & Battalio, R. C. (1980). Substitutability in time allocation. *Psychological Review*, **87**, 355-374.
- Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical procedures: real-time and retrospective evaluations of two minimally invasive treatments. *Pain*, **66**, 3-8.
- Roberts, W. W. (1958). Both rewarding and punishing effects from stimulation of posterior hypothalamus of cat with same electrode at same intensity. *Journal of Comparative and Physiological Psychology*, **51**, 400-407.
- Routtenberg, A., & Lindy, J. (1965). Effects of the availability of rewarding septal and hypothalamic stimulation on bar pressing for food under conditions of deprivation. *Journal of Comparative and Physiological Psychology*, **60**(2), 158-161.
- Sax, L., & Gallistel, C. R. (1990). Characteristics of spatiotemporal integration in the priming and rewarding effects of medial forebrain bundle stimulation. *Behavioral Neuroscience*, **105**, 884-900.
- Schreiber, C. A., & Kahneman, D. (submitted). Determinants of the remembered utility of aversive sounds. .
- Shizgal, P. (1997). Neural basis of utility estimation. *Current Opinion in Neurobiology*, **7**(2), 198-208.
- Shizgal, P., & Conover, K. (1996). On the neural computation of utility. *Current Directions in Psychological Science*, **5**(2), 37-43.
- Shizgal, P., Conover, K., & Arvanitogiannis, A. (1996). Performance for brain stimulation reward as a function of the rate and magnitude of reinforcement. *Society for Neuroscience Abstracts*, **22**(1), 686.

- Shizgal, P., & Matthews, G. (1977). Electrical stimulation of the rat diencephalon: Differential effects of interrupted stimulation on on- and off-responding. *Brain Research*, **129**, 319-333.
- Shizgal, P., & Murray, B. (1989). Neuronal basis of intracranial self-stimulation. In J. M. Liebman & S. J. Cooper (Eds.), *The neuropharmacological basis of reward* (pp. 106-163). Oxford: Oxford University Press.
- Simmons, J. M., & Gallistel, C. R. (1994). Saturation of subjective reward magnitude as a function of current and pulse frequency. *Behavioral Neuroscience*, **108**, 151-160.
- Stephens, D. W., & Krebs, J. R. (1986). *Foraging Theory*. Princeton, NJ: Princeton University Press.
- Treue, S., & Maunsell, J. H. R. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*, **382**, 539-541.
- Wasserman, E. M., Gomita, Y., & Gallistel, C. R. (1982). Pimozide blocks reinforcement but not priming from MFB stimulation in the rat. *Pharmacology Biochemistry & Behavior*, **17**, 783-787.
- Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature*, **382**, 629-632.
- Williams, B. A. (1988). Reinforcement, choice, and response strength. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology: Learning and cognition* (2nd ed., Vol. 2, pp. 167-244). New York: Wiley.
- Wise, R. A. (1996). Addictive drugs and brain stimulation reward. *Annual Review of Neuroscience*, **19**, 319-340.
- Yeomans, J. S. (1988). Mechanisms of brain-stimulation reward. In A. E. Epstein & A. R. Morrison (Eds.), *Progress in Psychobiology and Physiological Psychology* (Vol. 13, pp. 227-265). New York: Academic Press.
- Yeomans, J. S. (1990). *Principles of Brain Stimulation*. New York: Oxford University Press.

Young, P. T. (1967). Palatability: the hedonic response to foodstuffs. In C. o. d. e. CF, W. Heidel, J. R. Brobeck, & R. K. Crane (Eds.), *Handbook of physiology*. (Vol. 1, pp. 353-366). Washington, D.C.: American Physiological Society.

Zajonc, R. B. (1980). Feeling and thinking: preferences need no inference. *American Psychologist*, **35**(2), 151-175.

Counter Model

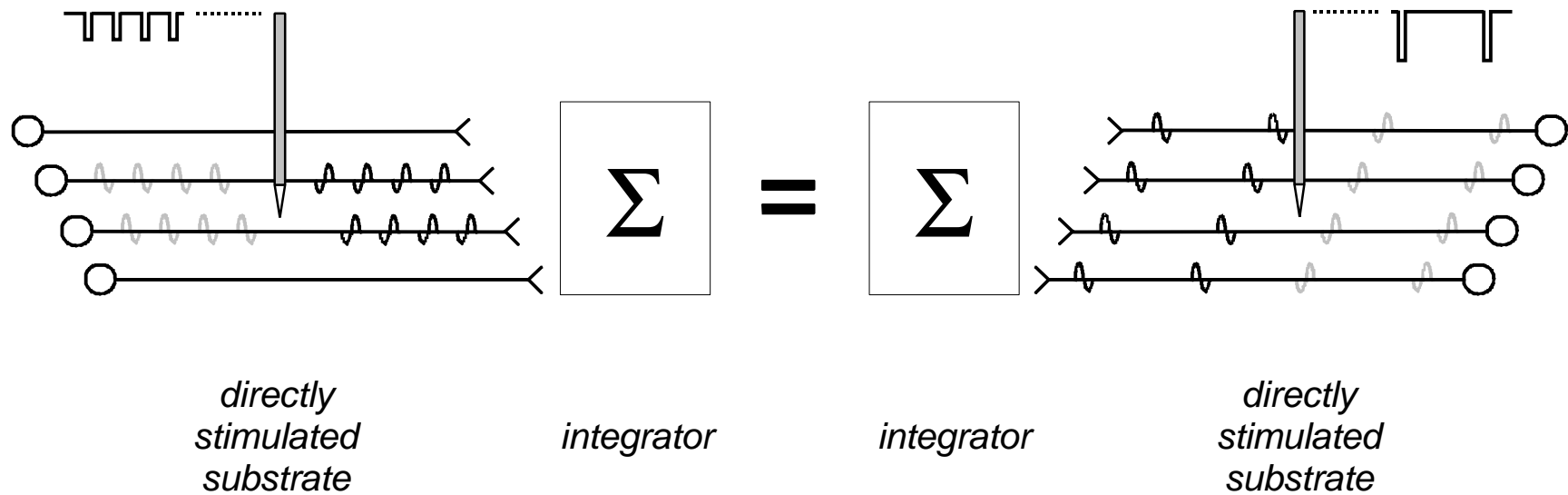


Figure 1

BSR: computing instantaneous utility

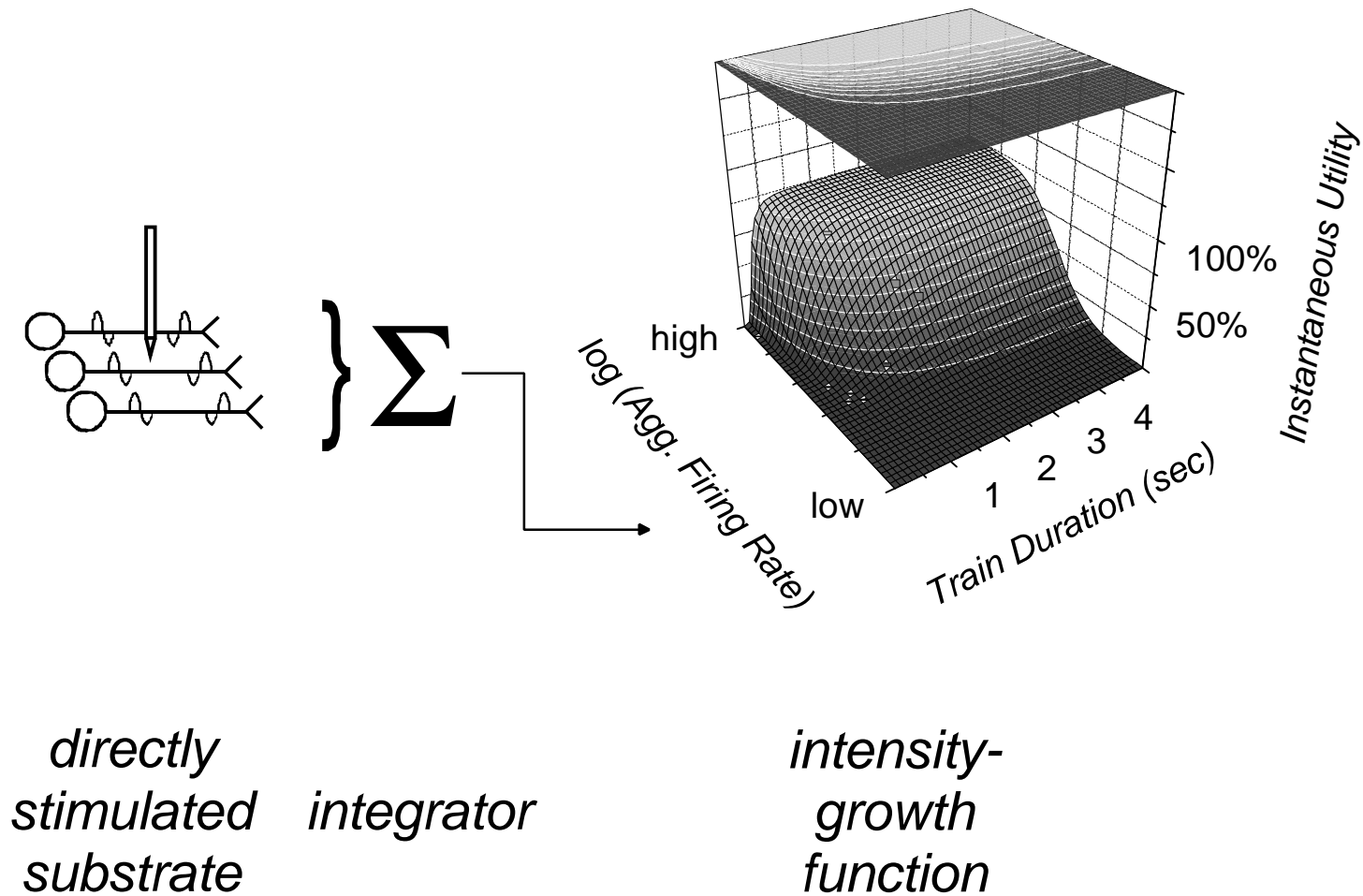


Figure 2

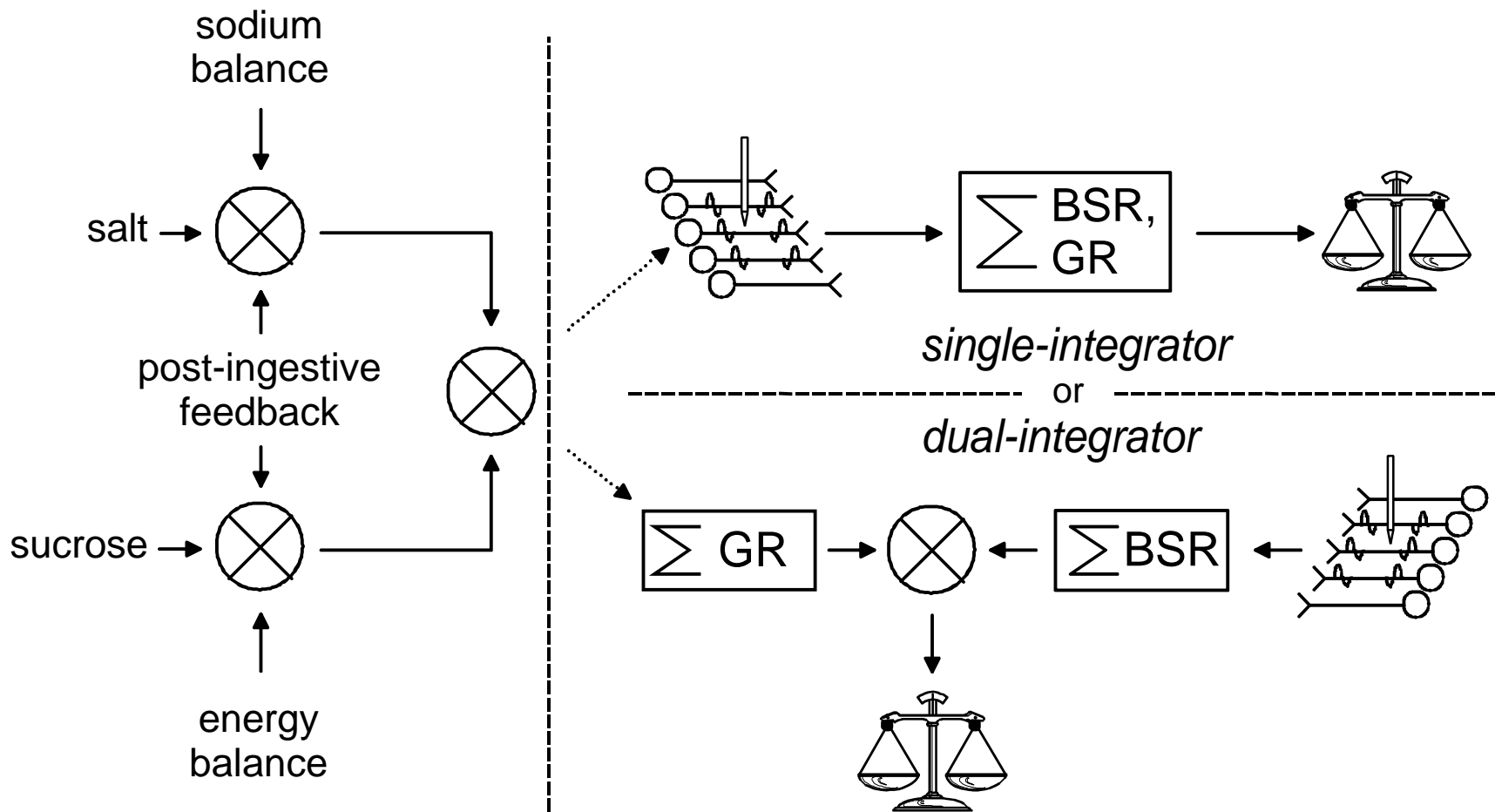


Figure 3

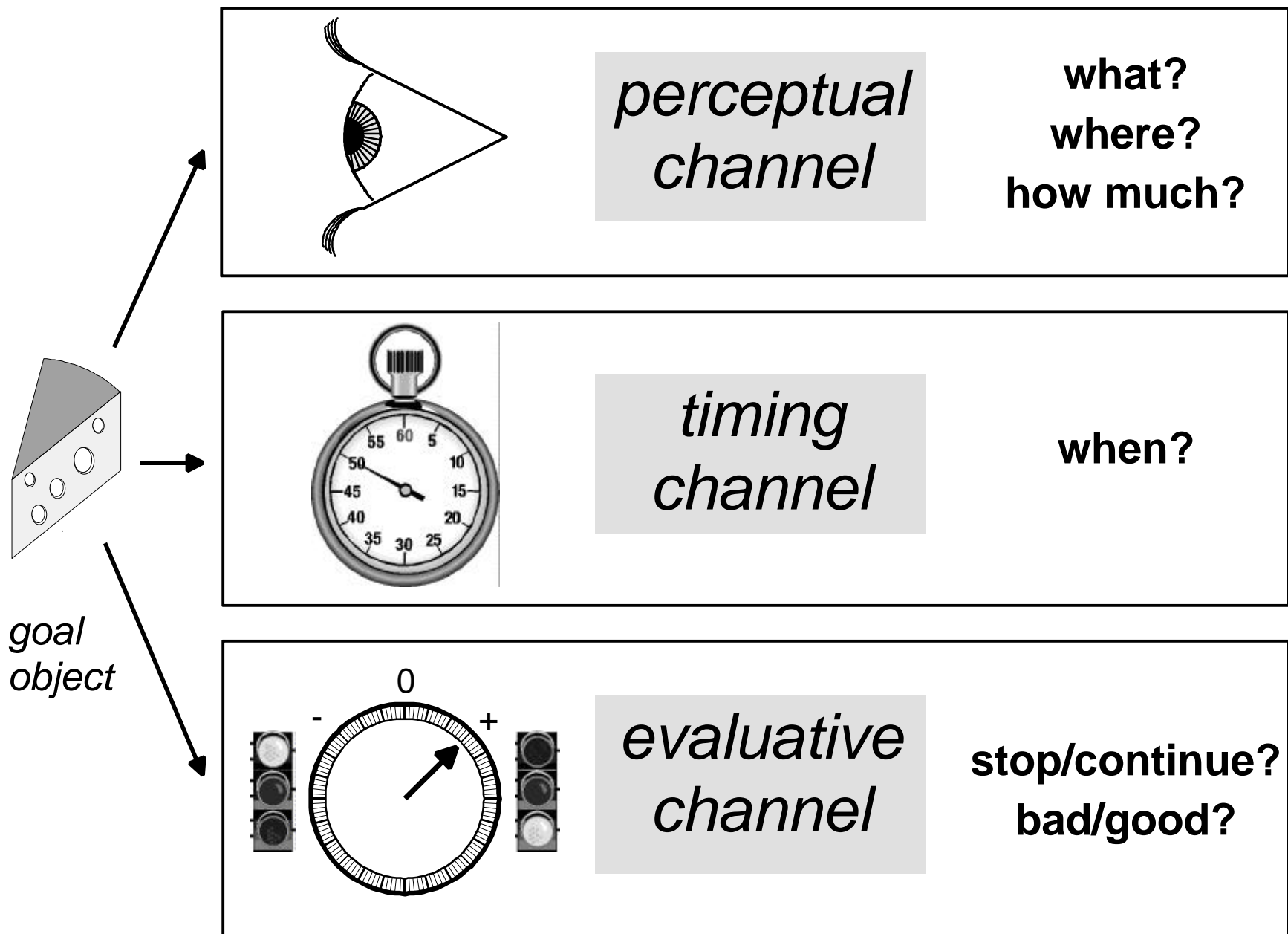


Figure 4

BSR: computing subjective intensity

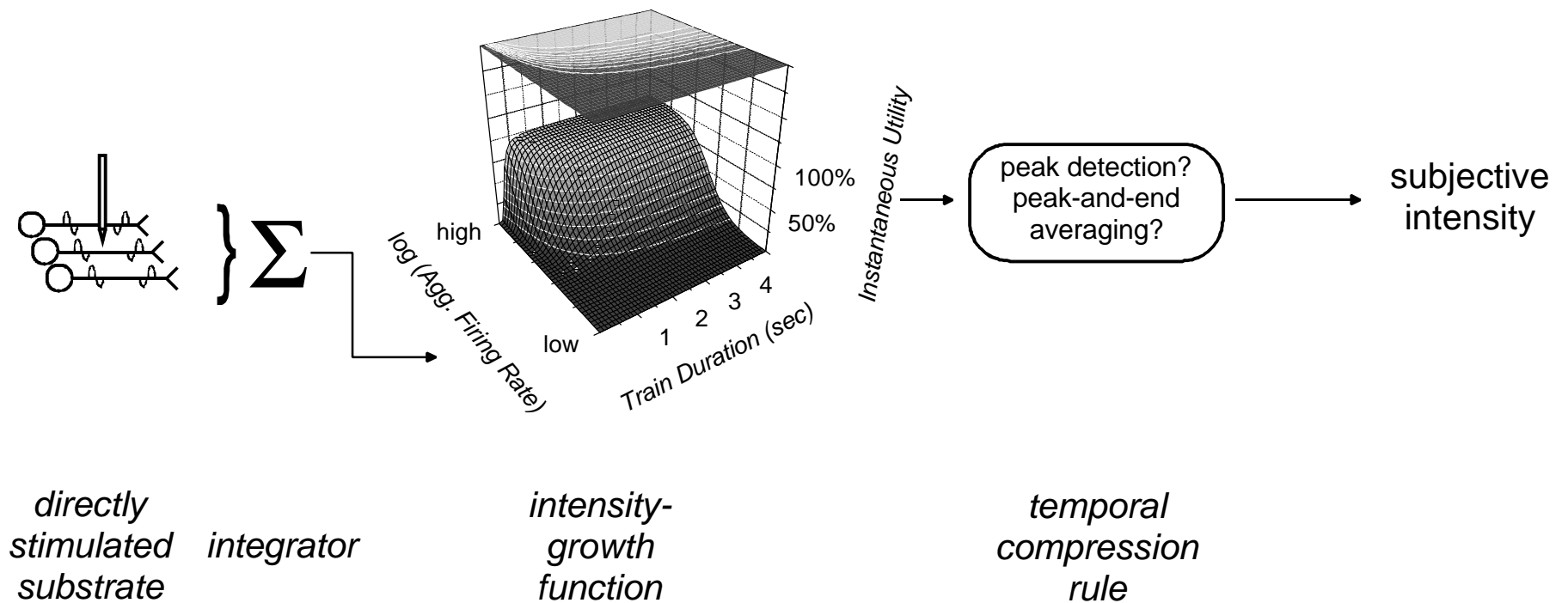


Figure 5

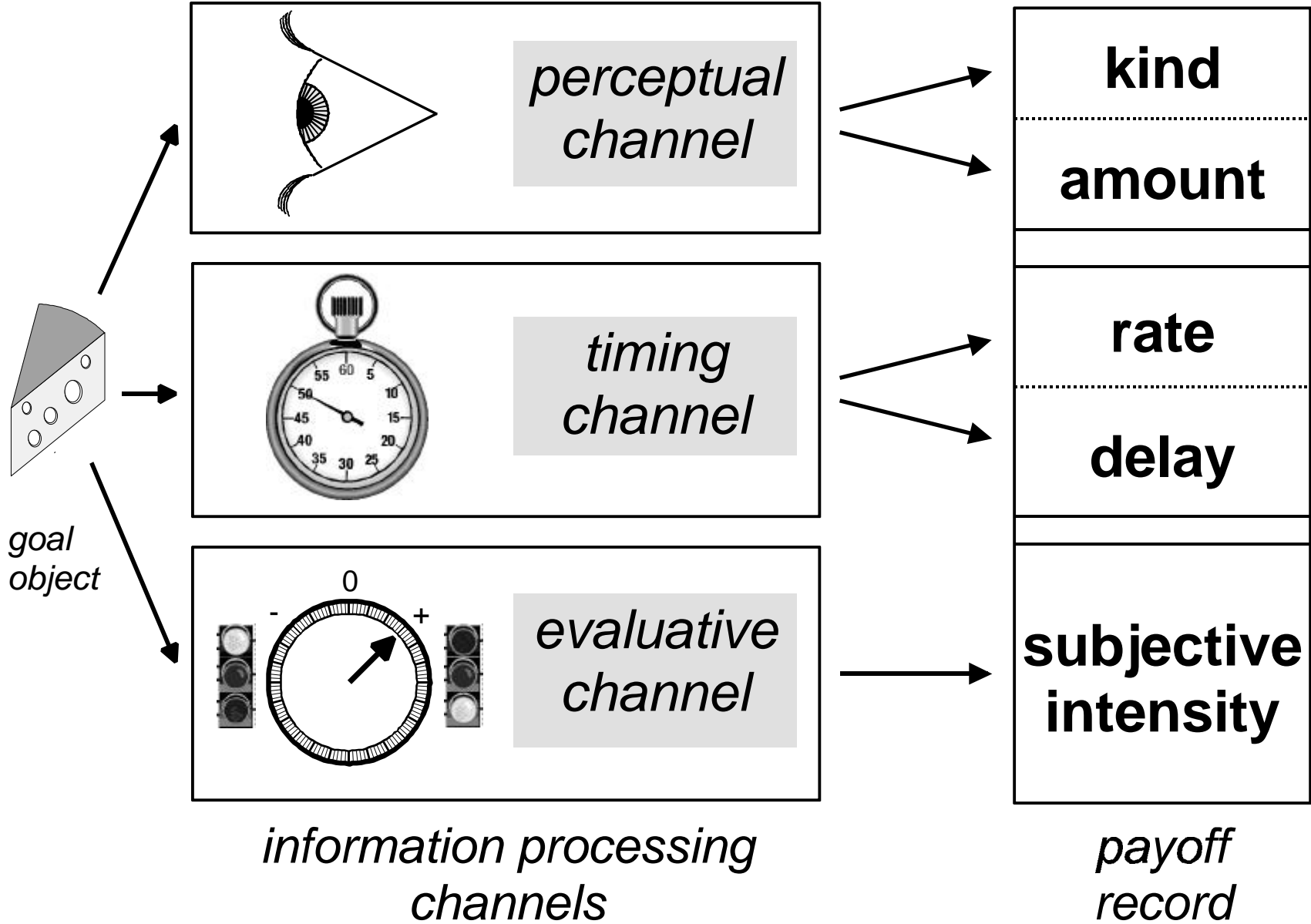


Figure 6

BSR: computing subjective payoff

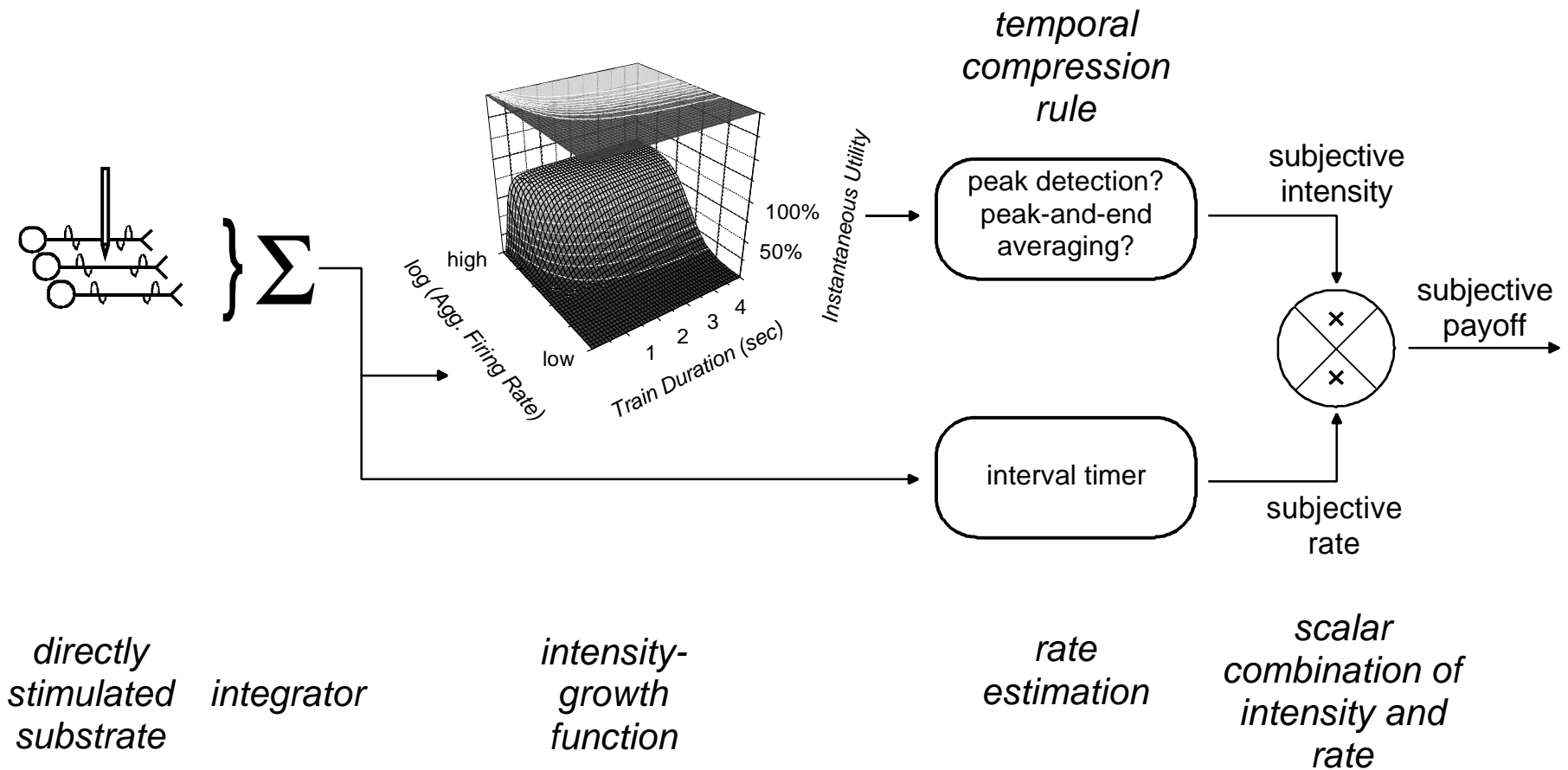


Figure 7

Computing Behavioral Allocation

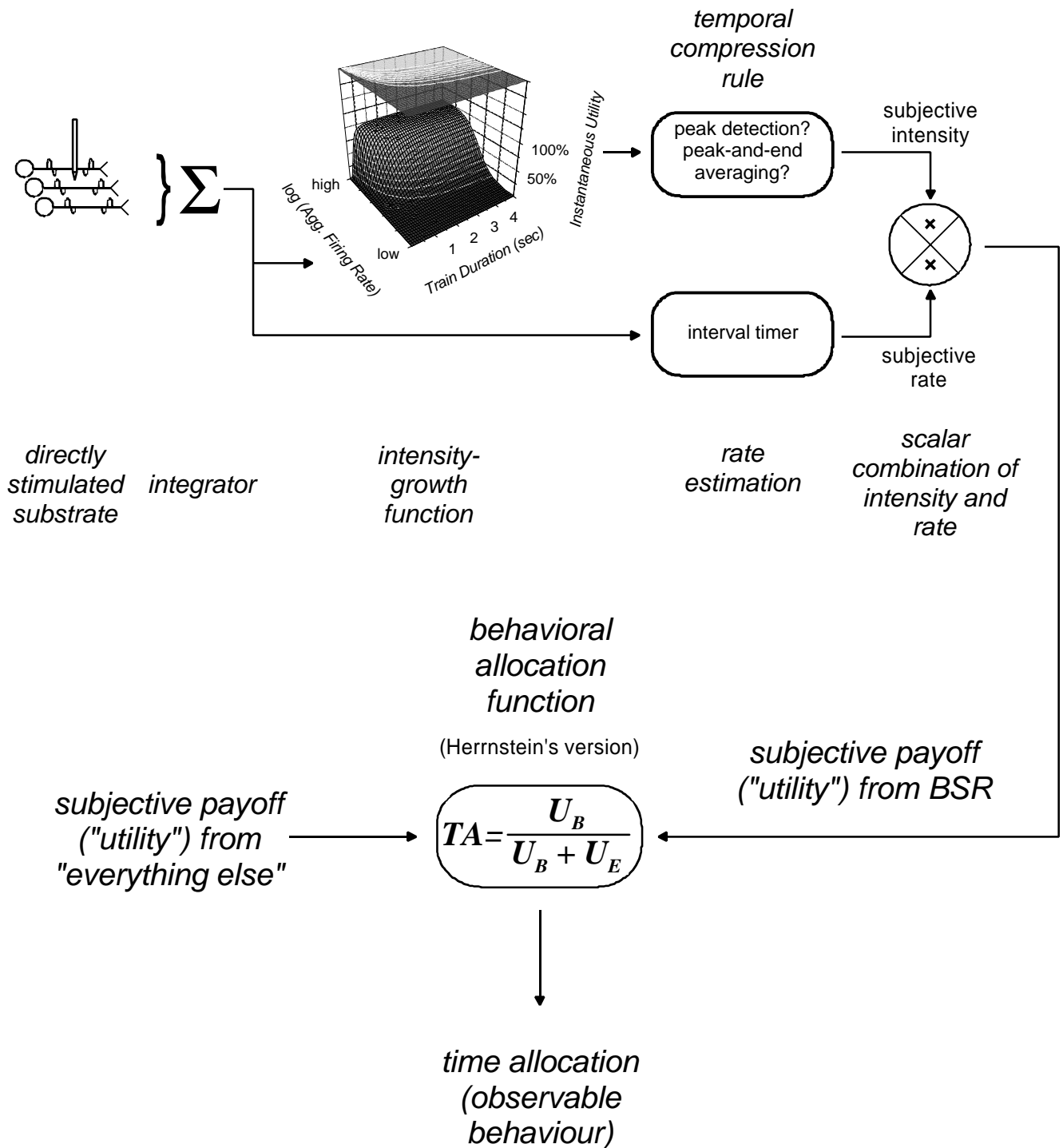


Figure 8