Efficient decision-makers evaluate relative reward per effort

Jan Kubanek[1†]

[1] University of Utah

***For correspondence:** jan.kubanek@utah.edu

Present address: [†] University of Utah, Salt Lake City, UT, United States

Abstract Understanding how humans and animals can make effective decisions given natural constraints would have a profound impact on economics, psychology, ecology, and related fields. Neoclassical economics provides a formalism for optimal decisions, but the apparatus requires a large number of evaluations of the decision options as well as neural representations and computations that are not biologically plausible. This article shows that specific biological considerations distill the economic optimization into an efficient and biologically plausible decision strategy. In this strategy, decision-makers compare the relative reward per relative effort of each option, and allocate their effort proportionally. Using a combination of analytical and simulation approaches, the study shows that this strategy is efficient, providing optimal or near-optimal gain following a single evaluation of decision options. The strategy is also rational; greedy and indifferent decision-makers are found to perform relatively poorly. Moreover, the relativistic value functions underlying this efficient strategy provide an account of time discounting, effort discounting, and resource discounting in general.

Introduction

Neoclassical economics provides a powerful apparatus that dictates how humans and animals should allocate their time, effort, or monetary resources (***Samuelson, 1948***; ***Rachlin et al., 1981***). A hallmark of this normative set of theories is that decision-makers maximize a criterion—individuals allocate their resources to maximize their total utility or payoff (***Friedman, 1953***; ***Becker, 1978***; ***Binmore, 1991***).

Albeit normative, this set of theories has faced major criticisms (***McNamara and Houston, 1980***; ***Hutchinson and Gigerenzer, 2005***; ***Stephens and Krebs, 2019***). The most severe is that there is no specific prescription for how individuals should make decisions to maximize their total utility (***Karatzas and Wang, 2000***). To attain the prescribed maximum, it is assumed that decision-makers can represent and perform calculus in multidimensional space and evaluate decision option in a near-limitless manner (***Simon, 1955***; ***Hollis and Nell, 1975***; ***Vaughan Jr and Herrnstein, 1987***). Moreover, the apparatus requires an absolute definition of utility (***Simon, 1956***; ***Robinson, 1962***).

To address these biologically implausible requirements, this study incorporates into the nor-

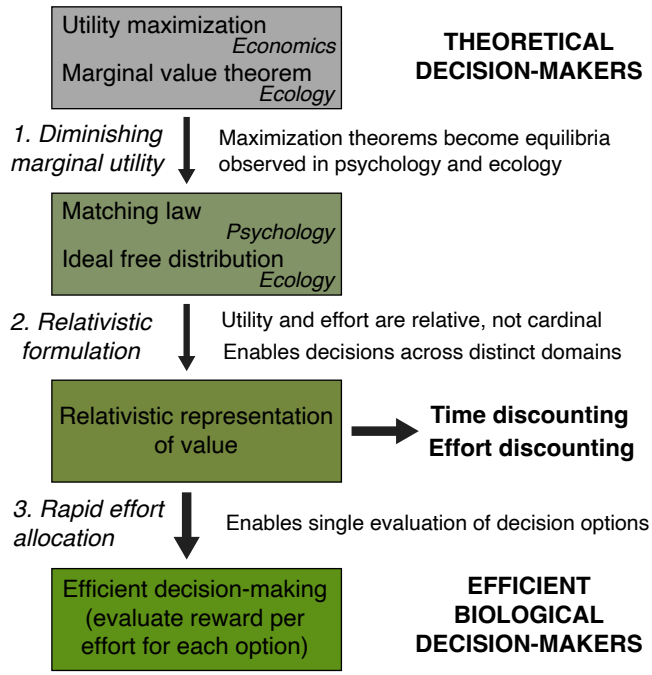

Figure 1. Framework for efficient biological decision-making. Steps, detailed formally below, that relate the maximization framework of economics (gray, top box) to value representations in psychology and an efficient biological decision framework (green, bottom box). These steps take into account natural and biological phenomena and constraints (see text for details).

mative framework specific biological considerations. These considerations distill this normative framework into a simple yet powerful strategy that provides efficient and rational decisions that can be performed by biological decision-makers.

Results

From economic maximization to efficient biological decision-making

The economic maximization framework provides an analytical prescription for the highest possible gain given resource constraints (***Rachlin et al., 1981***), and thus constitutes an essential starting point (Figure 1, top). According to this framework, subjects maximize the total expected utility across their n options, $\sum_{i=1}^{n} U_i(E_i)$. Here, U_i is the expected utility of an option and E_i the effort or expenditure required to produce that utility. The framework takes into account the decision-maker's limited resources, such as effort or time, $\sum_{i=1}^{n} E_i = E_{\max}$. (The terms expenditure, effort, resource, and behavior are used interchangeably in this text as individuals' means to obtain an outcome, and are all labeled E.) The maximization of total utility requires that effort or resources be allocated to the options such that the marginal utilities across all options are equal (***Kubanek, 2017***):

$$\frac{dU_1(E_1^*)}{dE_1} = \frac{dU_2(E_2^*)}{dE_2} = \ldots = \frac{dU_n(E_n^*)}{dE_n}. \tag{1}$$

When effort embodies foraging time ($E = T$), an equivalent form of this equation has been presented as the *marginal value theorem* in ecology (***Charnov, 1976***).

Step 1. Apply the law of diminishing marginal utility.

A previous study (***Kubanek, 2017***) showed that the marginal fractions of Equation 1 can be expressed as absolute fractions for an infinite number of $U(E)$ relationships that follow the law of diminishing marginal utility (***Friedman and Savage, 1948***). This leads to the following simplification:

$$\frac{U_1(E_1^*)}{E_1^*} = \frac{U_2(E_2^*)}{E_2^*} = \ldots = \frac{U_n(E_n^*)}{E_n^*}. \tag{2}$$

In psychology, this commonly observed equilibrium state is known as the *matching law* (***Herrnstein, 1961***; ***Baum, 1974***; ***Herrnstein, 2000***; ***Houston et al., 2021***). In ecology, the expenditure of a society of foraging individuals is equivalent to the number of individuals N_i devoted to each foraging patch. In this case, $E_i^* = N_i$, and the equation becomes the *ideal free distribution* (***Kacelnik et al., 1992***; ***Houston et al., 2021***), $\frac{U_1}{N_1} = \frac{U_2}{N_2} = \ldots = \frac{U_n}{N_n}$.

Step 2. Cast utility and effort in relative terms.

Albeit simpler, Equation 2 requires decision-makers to enumerate cardinal, numeric utilities. To address this issues, observe that each of the fractions in Equation 2 amounts to the same value. Let us denote this value as k. Accordingly, $U_i(E_i^*) = kE_i^*$, and so $\frac{\sum_i U_i(E_i^*)}{\sum_i E_i^*} = k$. Therefore,

$$\frac{U_1(E_1^*)}{E_1^*} = \frac{U_2(E_2^*)}{E_2^*} = \ldots = \frac{U_n(E_n^*)}{E_n^*} = k = \frac{\sum_i U_i(E_i^*)}{\sum_i E_i^*}. \tag{3}$$

The normalization by k leads to a relativistic form of the matching law:

$$\frac{U_1(E_1^*)}{kE_1^*} = \frac{U_2(E_2^*)}{kE_2^*} = \ldots = \frac{U_n(E_n^*)}{kE_n^*} = 1, \tag{4}$$

where $k = \frac{\sum_i U_i}{\sum_i E_i^*}$ (see Equation 3).

The normalized fractions in this equation are prime candidates for a rational representation of value. First, the expressions include both the benefits (reward) and costs (effort) associated with each option, in line with the axioms of rational choice theory (***Buchanan, 1978***). Second, the values can be compared across options and are equalized in equilibrium (Equation 4). Indeed, the value representation $V = \frac{U}{kE}$ captures several phenomena in psychology (Figure 1). In particular, decisions that involve time investment ($E = T$) and yield monetary or other outcomes A are characterized with $V = \frac{A}{kT}$. This expression embodies *time discounting* (***Myerson and Green, 1995***; ***Louie and Glimcher, 2010***). Moreover, many situations require an investment of physical or mental effort E to obtain a reward. In these cases, the value function takes the form $V = \frac{A}{kE}$. This embodies *effort discounting* (***Białaszek et al., 2019***; ***Botvinick et al., 2009***). The Materials and Methods section proves that value functions in the form $V = \frac{U}{kE}$ are tied to economic maximization under the law of diminishing marginal utility in a way analogous to the previously demonstrated $V = \frac{U}{E}$ (***Kubanek, 2017***).

Critically, the normalization of Equation 3 by $k = \frac{\sum_i U_i}{\sum_i E_i^*}$ turns absolute utilities U_i into relative utilities $u_i = \frac{U_i}{\sum_k U_k}$. This relativistic formulation circumvents the problematic cardinal definition of utility (***Simon, 1956***; ***Robinson, 1962***), and is in this text referred to as "reward". Furthermore, the

normalization turns the absolute effort values E_i into relative efforts $e_i = \frac{E_i}{\sum_k E_k}$. This leads to the following result:

$$\frac{u_1(e_1^*)}{e_1^*} = \frac{u_2(e_2^*)}{e_2^*} = \ldots = \frac{u_n(e_n^*)}{e_n^*} = 1. \tag{5}$$

This result states that decision-makers should distribute relative efforts in proportion to relative rewards of their options:

$$e_i^* = u_i(e_i^*). \tag{6}$$

This relativistic formulation enables decision-makers to make choices across options of distinct kinds (see Discussion), and is crucial for the final step.

Step 3. Derive an effort allocation strategy based on the relativistic formulation.

Equation 5 embodies a relativistic form of the matching law, which represents an equilibrium state. This equilibrium is rational under the law of diminishing marginal utility (***Kubanek, 2017***). This section turns the formalism into a time-by-time strategy that enables humans and animals to make efficient decisions. The relativistic formulation of the previous section is key to this step. In this formulation (Equation 6), both sides of the equation, e_i and u_i, are on the same relativistic scale and $\in (0, 1)$. This way, the ideal $e_i^* = u_i(e_i^*)$ can be approximated through an iterative rule

$$e_i^{\text{next}} = u_i(e_i^{\text{prev}}). \tag{7}$$

For a single evaluation of the decision options, this rule becomes

$$e_i^{\text{harvest}} = u_i(e_i^{\text{eval}}), \tag{8}$$

where e_i^{eval} is the initial assessment of the worth of the decision options, and e_i^{harvest} is the resulting effort allocation.

In what follows, I demonstrate that a single application of Equation 7 (i.e., Equation 8) provides optimal or near-optimal gain. Specifically, the Materials and Methods section proves that a single evaluation of the decision options provides optimal gain when the reward functions u_i saturate logarithmically with respect to effort. Furthermore, the Materials and Methods section shows that a single evaluation also converges to an optimum for exponential u_i that saturate strongly with increasing effort. Finally, simulations performed below show that when this effort allocation strategy is applied repeatedly (Equation 7), it provides high performance even in situations in which the law of diminishing returns does not hold.

Given the high performance yet rapid and simple evaluation that can be performed by biological decision-makers, the effort allocation strategy characterized by Equation 7 and Equation 8 is referred to as "Efficient Decision Making" (EDM). I compare the efficiency of EDM with that of the theoretical economic agents and evaluate its rationality in regard to its greedy and indifferent counterparts.

Performance of Efficient Decision Making

The analytical findings of EDM efficiency are herein validated through simulations. Performance is evaluated over decision situations that include 2, 3, and up to 10 options, each with distinct, randomly generated $U_i(E_i)$ relationships, for a total of 9000 choice situations (see Materials and Methods).

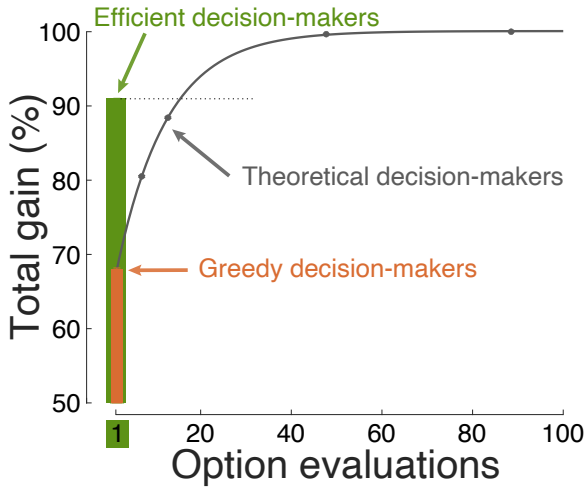

Figure 2. Efficient decision making following a single evaluation of decision options. The figure compares the performance of three kinds of decision-makers. The first kind (gray) performs multidimensional optimization over the decision options (see Materials and Methods). The second kind (green) uses efficient decision making. The third kind (orange) uses a greedy strategy, putting all resources into the most desirable option. The decision-makers faced a total of 9000 simulated choice situations, involving up to 10 options, each with distinct contingencies between utility and effort (Materials and Methods). The maximum gain (the value of 100%) was established by allowing the multidimensional optimization algorithm to converge. EDM (green) provides a high gain while requiring a single evaluation step. Here and henceforth, all plots show mean±s.e.m. values. The error bars are often smaller than the symbols indicating the mean.

Efficiency of EDM

The efficiency of EDM is evaluated by comparing the total gain of decision-makers using this strategy ("efficient decision-makers") with the maximal possible gain of theoretical economic agents (see Materials and Methods for details). Figure 2 shows the result. Efficient decision-makers obtained an average 91.0% of the total possible gain across the choice situations tested. Crucially, they collected this level of gain following a single evaluation of their options. In contrast, the theoretical economic agents, whose computations involve multidimensional optimization, require an average of at least 15 evaluations of the decision options to reach the same level of performance (Figure 2, gray). To put this into natural perspective, consider a foraging animal that must decide how to allocate her presence across several patches of food. If she was (and could) use the economic framework (Equation 1), she would have to visit and forage in each patch at least 15 times to make a decision of quality comparable to EDM. EDM allows biological decision-makers to make effective decisions in a single step, i.e., efficient decisions.

Rationality of EDM

The rationality of EDM is evaluated with respect to the performance of greedy and indifferent decision-makers. A greedy decision-maker does not perform the reward per effort valuation of Equation 5, and instead puts all effort or resources into the option that provides the highest reward. Figure 2 (orange) shows that the greedy approach severely degrades performance. On average, greedy decision-makers only harvested an average of 68.0% of the total possible gain across their options. The difference between the rational and the greedy decision-makers is substantial and

highly significant ($t_{8999} = 110.26$, $p \doteq 0$, two-sided t-test).

In EDM, relative effort is proportional to relative reward, i.e., $e_i^{\text{harvest}} = g u_i(e_i^{\text{eval}})$ with $g = 1.0$. Varying this proportionality constant enables us test the performance of decision-makers of varying levels of greed ($g > 1.0$) as well as individuals who are relatively indifferent to reward ($g < 1.0$). Figure 3a shows that $g = 1.0$ indeed lies near the optimum, with $g = 1.0$ and $g = 0.8$ both attaining an average 91% gain. In comparison, the greedy and indifferent strategies, albeit equally rapid, are relatively ineffective.

EDM can be applied in more than one evaluation step to increase the total gain further. To do so, the allocation $e_i^{\text{next}} = u_i(e_i^{\text{prev}})$ is repeated a desired number of times. Figure 3b, left green bars, shows that a second evaluation increases the average performance from 91.0% to 92.4%. The performance saturates in about the third and fourth evaluation, reaching 92.6% of the possible maximum. In comparison, the performance of the greedy strategy does not, by definition, improve with additional evaluations (Figure 3b, left orange bars).

Thus far, we have assumed that decision-makers allocate effort or expenditures equally when sampling the initial worth of their options (Materials and Methods). However, equal sampling may be impractical or impossible in many choice situations. The following analysis reveals that EDM is robust with respect to a particular initial distribution e_i^{eval}. Specifically, Figure 3b, right green bars, shows that decision-makers who distribute their initial effort e_i^{eval} entirely randomly can still apply the same strategy with comparable level of efficiency. In addition, the performance increases even more steeply by re-adjusting effort in response to additional option evaluations ($e_i^{\text{next}} = u_i(e_i^{\text{prev}})$). On the fourth evaluation, the performance reaches 92.1%. A two-way ANOVA with factors evaluation number (I) and initial effort distribution (D) detected a significant effect of I ($F(3, 71992) = 502.84$, $p \doteq 0$), a significant effect of D ($F(1, 71992) = 526.95$, $p = 3.4 \times 10^{-116}$), and a significant interaction between the factors ($F(3, 71992) = 42.35$, $p = 2.5 \times 10^{-27}$). The greedy strategy does not depend on the number of evaluations (Figure 3b, orange) but is sensitive to the initial allocation of effort ($F(1, 71992) = 343.03$, $p = 2.1 \times 10^{-76}$).

EDM is robust also with respect to the number of decision options. Figure 3c shows that decision-makers using EDM perform well even when faced with a relatively high number of options. EDM attains an average 93.5% gain for all situations that involved 2 options. The strategy has an exponential decay of 2.3% per option, and achieves a 87% gain for 10-option decisions. In contrast, greedy decision-makers fall short when facing an appreciable number of options. This is expected because putting all effort into the most rewarding option leads to an opportunity loss. The greedy strategy shows an exponential decay of 3.9% per option, and attains a mere 52.8% average gain in 10-option situations.

The law of diminishing marginal utility, which was harnessed to derive EDM, may not hold in all situations. In certain situations, a payoff may increase indefinitely with the invested effort. To investigate the effect of this possibility on EDM, let us parameterize the strength of diminishing returns using the power function, $U(E) = U_m E^{\alpha}$, where α is the tested parameter. For $\alpha = 0.1$, the function exhibits strongly diminishing returns. On the other hand, for $\alpha = 1.0$, gain increases linearly with effort and never saturates. Figure 3d shows the dependence of the total gain on this parameter. As expected, EDM, based on the law of diminishing marginal utility, shows high performance when

the choice options obey the law. The greedy strategy is particularly inadequate in this case. Greed is blind to the law of diminishing marginal utility: A greedy decision-maker puts all effort or resources into an option whose payoff eventually saturates, thus presenting an opportunity loss. As expected, the performance of EDM / the greedy strategy is lower / higher in choice situations in which the law of diminishing marginal utility does not hold ($\alpha = 1.0$). Nonetheless, EDM can outperform the greedy strategy in such cases, when applied iteratively. In particular, following the fourth evaluation, the efficient strategy attains a 91.2% gain, compared to a 88.8% gain of the greedy approach. The same principal observation is made for functions other than the power function (Figure 4).

Discussion

This study derives an efficient decision strategy from the normative economic framework. The key finding is that efficient decisions are made by evaluating the relative reward per relative effort associated with the decision options (EDM strategy). EDM is efficient in that it provides optimal or near-optimal gain following a single evaluation of the decision options. EDM is also rational in that its greedy and indifferent counterparts perform relatively poorly.

EDM is simple yet highly efficient. In particular, EDM was benchmarked across 9000 decision situations of up to 10 decision options, and harvested 91.0% of the maximum possible gain following a single evaluation of the decision options. Moreover, in human settings, in which the worth of the options may be known or can be estimated (e.g., price), EDM allows individuals to allocate effort (e.g., money) readily, without any prior sampling. Thus, EDM is efficient, combining high performance with rapid and simple evaluation.

A form of EDM has been considered by ecologists as a candidate for an evolutionarily stable strategy (***Harley, 1981***; ***Tracy and Seaman Jr, 1995***). Referred to as the *relative payoff sum* in behavioral ecology, a population using this strategy is believed to be uninvadable by a mutant with a different strategy (***Harley, 1981***; ***Hamblin and Giraldeau, 2009***). A proof was attempted (***Harley, 1981***) but remains elusive (***Houston and Sumida, 1987***). The present study shows that this strategy stems from the maximization framework of neoclassical economics, and shows an analytical relationship of EDM to this normative framework. Critically, analytical derivations and simulations show that the strategy is efficient, providing optimal or near-optimal gain in a single evaluation of the choice options in situations in which the law of diminishing marginal utility holds (Figure 3, Figure 4). Given that foraging efficiency is a critical optimization variable in ecology (***Charnov, 1976***), this study supports the notion of this strategy as evolutionarily stable, under the assumption that the law of diminishing marginal utility is a prevalent natural phenomenon.

Evidence for EDM has also been found in neural activity in the primate brain. Specifically, the relative value associated with EDM, termed “fractional income,” captures firing rates of neurons in monkey area LIP (***Sugrue et al., 2004***). The effect was specifically due to this local strategy; the matching law equilibrium had no leverage on the neuronal activity (***Sugrue et al., 2005***).

The value representations underlying EDM align with the empirically established value functions in psychology, including time and effort discounting. This study provides a teleological reason for such representations—efficiency with respect to limited resources, including time and effort, and their saturating returns. As a consequence, efficient decision-makers evaluate the reward rate (i.e.,

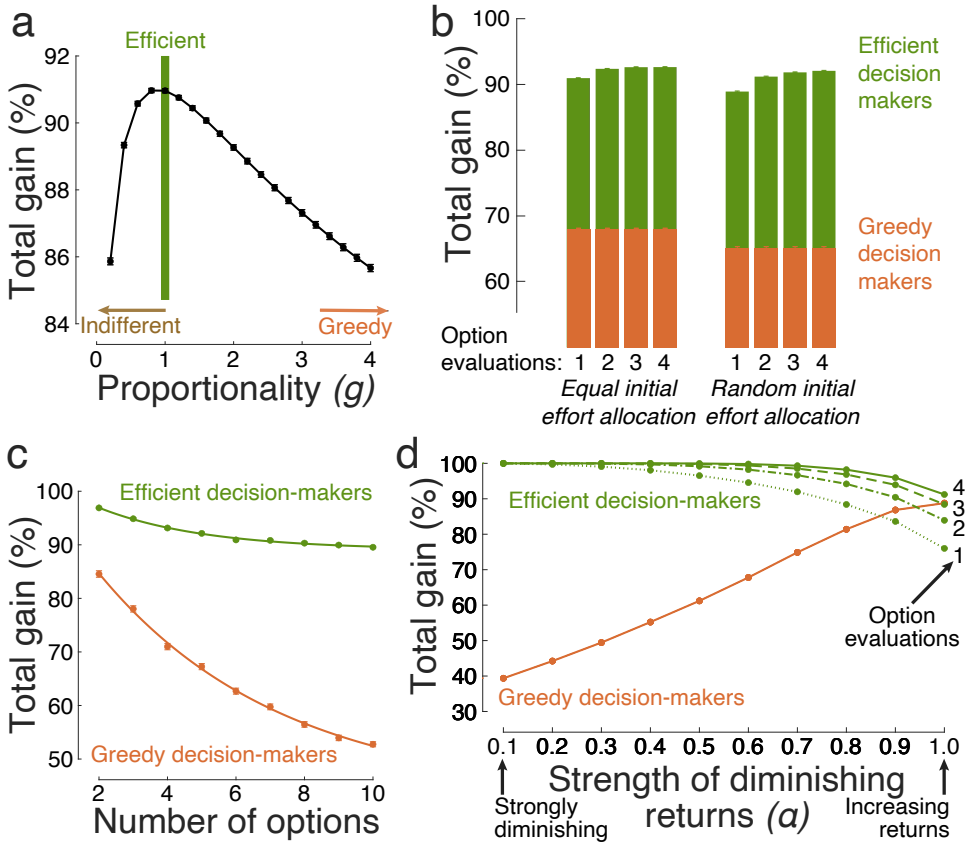

Figure 3. The performance and robustness of efficient decision making.
(**a**) Gain as a function of the level of greed g in $e_i^{\text{harvest}} = g u_i(e_i^{\text{eval}})$. The efficient, efficient decision making strategy uses $g = 1$, i.e, relative effort is allocated in proportion to relative reward. Decision-makers who are indifferent to reward ($g < 1$) and decision-makers who are greedy ($g > 1$) suffer a performance loss. Here and elsewhere in this figure, the same 9000 decision situations as in Figure 2 were used. (**b**) Decision-makers must evaluate decision options at least once to estimate their initial worth. In this initial evaluation, decision-makers may distribute their effort across the options equally (left bars). However, such distribution, even though desirable (see Materials and Methods) may not always be feasible. The right bars show that distributing initial effort across options entirely randomly provides a comparable level of performance for EDM. This more realistic, random initial distribution is used henceforth (c and d). In both cases, the options can be evaluated iteratively, i.e., $e_i^{\text{next}} = u_i(e_i^{\text{prev}})$, to increase performance further. (**c**) Same data as in Figure 2, separately for a specific number of decision options. The curves represent exponential fits. (**d**) Performance as a function of the level of diminishing marginal utility. To evaluate the full spectrum from strongly diminishing to not at all diminishing (linearly increasing), this analysis specifically used the power contingency $U(E) = U_m E^{\alpha}$. The abscissa shows performance as a function of 10 evenly spaced values of α (see Materials and Methods). Like in b, the decision strategy $e_i^{\text{next}} = u_i(e_i^{\text{prev}})$ can be applied iteratively. Performance curves for up to 4 iterations are shown in distinct line styles. The greedy strategy (orange), by definition, does not improve with additional iterations.

reward per time), the reward per effort, or, most generally, the reward per the applicable resource across their options. An example of the reward rate valuation are decisions concerning job choice based on salary (i.e., monetary income per unit of time).

Decision-makers are known to weigh benefits against costs, or rewards against effort. Yet, how effort should be incorporated in current models of decision-making in economics (***Buchanan, 1978***), ecology (***Zipf, 1949***), and psychology (***Białaszek et al., 2019***; ***Botvinick et al., 2009***) has been unclear. This study shows that individuals can make efficient cost-benefit decisions by evaluating the relative reward u per relative effort e of each option, i.e., $\frac{u}{e}$. This formulation increases with the benefit of an option (u), decreases with the cost of an option (e), and the relativistic expression does not require any free parameter.

The derived equations provide a normative insight into the parameter k used in studies of time and effort discounting. Specifically, $k = \frac{\sum U_i}{\sum E_i}$ serves to convert absolute utilities and efforts into relative quantities. Since k represents the total possible gain per total available resource, it provides specific predictions regarding the rates of discounting for each individual and each decision domain. For example, consider a decision-maker who can devote a limited amount of time to a task. The lower the amount of available time, the higher the k, and so the more steeply the individual discounts a gain with respect to a time investment ($V = \frac{U}{kT}$). Reversely, if a resource such as time is abundant, a payoff is discounted relatively little. Thus, the discounting factor incorporates the scarcity or the abundance of each decision-maker's resources.

The $\frac{u}{e}$ formulation derives from the normative economic framework (Figure 1), is found to be efficient (Figure 2) and rational (Figure 3), and is testable. For instance, time and effort discounting experiments in humans support this form of the value function (***Myerson and Green, 1995***; ***Louie and Glimcher, 2010***; ***Białaszek et al., 2019***; ***Botvinick et al., 2009***), and the rapid, trial-wise EDM evaluation has been observed in neuroscience (***Sugrue et al., 2004***).

The alignment of the value functions associated with EDM with value functions empirically obtained in psychology provides further validation of the considerations used to arrive at EDM (Figure 1). First, decision-makers have limited resources in the form of time and physical and mental energy. Second, expenditures of these resources commonly reach saturation or diminishing returns (***Gossen, 1854***; ***Cobb and Douglas, 1928***; ***Stephens and Krebs, 1986***; ***Kacelnik et al., 1992***). And third, decisions involve relative rather than absolute representations.

With respect to the representation of utility (***Simon, 1955***, ***1956***; ***Robinson, 1962***; ***Hollis and Nell, 1975***; ***Vaughan Jr and Herrnstein, 1987***), this study finds that individuals can make efficient decisions by representing utilities and efforts relativistically (Equation 5). This has important practical implications. First, relativistic representations enable neurons in the brain to encode the worth of decision options through relative firing rates (***Sugrue et al., 2004***) or other relative means. Second, relative utilities allow individuals to make decisions across distinct domains. Specifically, relative utilities or rewards can be evaluated over comparable domains (e.g., the domain of foods, the domain of activities). Like relative utilities, relative efforts can also be evaluated over comparable domains (e.g., the domain of time, the domain of mental effort). The associated value representations, $V = \frac{u_i}{e_i}$, can thus be used for decisions across distinct domains. Actions are directed to options in any domain that provide high relative reward per relative effort.

Notably, the u_i and e_i variables can be subjective. So long as the representations are comparable by the brain, the $e_i = u_i$ strategy provides an efficient allocation of the decision-maker's resources.

Frequently, payoffs in nature are probabilistic or vary over time. Several rules can be used to keep track of such changing environments and thus estimate their relative worth u_i. The most compact are related to reinforcement learning, in which previous payoffs are discounted geometrically using a "learning rate." This approach has been applied in ecology (***Harley, 1981***; ***Hamblin and Giraldeau, 2009***), computer science (***Sutton and Barto, 1998***), and neuroscience (***Sugrue et al., 2004***; ***Corrado et al., 2005***). Moreover, decision-makers can adapt the learning rate to the speed of change or the level of stochasticity of specific decision situations (***Iigaya et al., 2019***).

In summary, this study provides a prescription for how decision-makers can make efficient decisions. Optimal or near-optimal decisions are achieved by evaluating relative reward per relative effort across choice alternatives. Such decisions are made rapidly, using a single evaluation of the decision options. Moreover, the underlying value representations align with several observations in psychology, ecology, and neuroscience. This biologically plausible framework is expected to provide an anchor for understanding the efficiency and rationality of choice behavior in behavioral economics, psychology, ecology, and related fields.

Materials and Methods

Analytical derivations

Optimality of the relativistic formulation of the matching law

This section derives the conditions for which the relativistic form of matching, $\frac{U_1(E_1^*)}{kE_1^*} = \frac{U_2(kE_2^*)}{E_2^*} = \ldots = \frac{U_n(kE_n^*)}{kE_n^*} = 1$, $k = \frac{\sum_i U_i}{\sum_i E_i^*}$, is optimal. The proof proceeds in a similar fashion as in (***Kubanek, 2017***). In particular, the relativistic formulation of the matching law (Equation 5; see also Equation 3), which now in addition to (***Kubanek, 2017***) includes k, is optimal when the following relationship holds for each decision option i:

$$\frac{dU_i(E_i)}{dE_i} = h\left(\frac{U_i(E_i)}{kE_i}\right),$$

where h is a strictly monotonic function. For simplicity, here, let us label $U_i(E_i)$ as $U(E)$ and E_i as E.

The derivative of this equation with respect to respect to effort is

$$\begin{aligned}\frac{d^2U(E)}{dE^2} &= h'\left(\frac{U(E)}{kE}\right)\left(\frac{\frac{dU(E)}{dE}kE - kU(E)}{k^2E^2}\right)\\ &= h'\left(\frac{U(E)}{kE}\right)\left(\frac{\frac{\frac{dU(E)}{dE}}{k} - \frac{\frac{U(E)}{E}}{E}}{k}\right).\end{aligned}$$

Using the above $\frac{dU(E)}{dE} = h\left(\frac{U(E)}{kE}\right)$, we obtain:

$$\frac{d^2U(E)}{dE^2} = h'\left(\frac{U(E)}{kE}\right)\left(\frac{\frac{h\left(\frac{U(E)}{kE}\right)}{k} - \frac{U(E)}{kE}}{E}\right).$$

As in the main text, let the term $\frac{U(E)}{kE}$ represent a value function V, $V = \frac{U(E)}{kE}$. The beauty of this representation shines through the above equation. Given that $E > 0$, it is apparent that the sign of

the right side of the equation is solely a function of V:

$$\frac{d^2U(E)}{dE^2} = h'(V)\left(\frac{\frac{1}{k}g(V) - V}{E}\right).$$

In order for a critical point to represent a reward maximum, it must be $\frac{d^2U(E)}{dE^2} < 0$ for at least one option and at least for some value of E (***Kubanek, 2017***). To meet that requirement, and taking into account $E > 0$, there are two possibilities:

1. $h'(V) > 0$ and $h(V) < kV$
2. $h'(V) < 0$ and $h(V) > kV$

The second possibility is inadmissible because for a general $V > 0$, a function $h(V)$ cannot be decreasing while maintaining a value above a certain line $h(V) = kV$, $k > 0$. This leaves us with the first set of requirements, $h'(V) > 0$ and $h(V) < kV$.

For relativistic matching, $V = \frac{U_1(E_1)}{kE_1} = \frac{U_2(E_2)}{kE_2} = \ldots = \frac{U_n(E_n)}{kE_n}$ is the same for all options i. Thus, when $\frac{d^2U_i(E_i)}{dE_i^2} = h'(V)\left(\frac{\frac{1}{k}h(V)-V}{E_i}\right)$ assumes a positive or a negative sign, it does so equally for all options, and the sign is governed solely by the properties of $h(V)$. For matching yielding a reward maximum ($h'(V) > 0$ and $h(V) < kV$), it is then apparent that $\frac{d^2U_i(E_i)}{dE_i^2} < 0$ for all options and for any E_i. Thus, all options yield diminishing returns for all values of effort.

Optimality of a single evaluation

This section derives a general set of functions for which efficient decision making is optimal in a single option evaluation step. Consider a general function $U_i(E_i) = U_{m_i} f(E_i)$. Equation 1 states that $\frac{dU_1(E_1^*)}{dE_1} = \frac{dU_2(E_2^*)}{dE_2} = \ldots = \frac{dU_n(E_n^*)}{dE_n}$. Thus, $\frac{dU_i(E_i^*)}{dE_i} = z$, where z is a constant, and

$$\frac{df(E_i^*)}{dE_i} = \frac{z}{U_{m_i}}. \tag{9}$$

Let efficient decision making be optimal in a single evaluation step, $e_i^* = u_i(e_i^{\text{eval}})$, and let the initial evaluation e_i^{eval} of the decision options be distributed equally. Using Equation 4, we can also state this fact as $E_i^* = \frac{1}{k}U_i(E_i^{\text{eval}}) = \frac{U_{m_i}}{k}f(E_i^{\text{eval}})$. From here, $U_{m_i} = \frac{E_i^* k}{f(E_i^{\text{eval}})}$. Inserting this expression into Equation 9, we obtain

$$\frac{df(E_i^*)}{dE_i} = \frac{zf(E_i^{\text{eval}})}{E_i^* k}.$$

For a general function argument E_i^*, a derivative proportional to $\frac{1}{E_i^*}$ can be obtained for logarithmic functions. Thus, functions in the general form $U(E) = U_m \log_a(\frac{E}{b}) + c$, where $U_m > 0$, $a > 1$, $b > 0$, and $c \in R$ are free parameters, provide the sought space of solutions.

Let us validate this result. When the decision-maker distributes her initial effort equally across the options, $E_1^{\text{eval}} = E_2^{\text{eval}} = \ldots = E_n^{\text{eval}} = E_{\text{equal}}$, she harvests, for a general function $U_i(E_i) = U_{m_i} f(E_i)$

$$u_i(e_i^{\text{eval}}) = \frac{U_i(E_i)}{\sum_k U_k(E_k)} = \frac{U_{m_i} f(E_{\text{equal}})}{\sum_k U_{m_k} f(E_{\text{equal}})} = \frac{U_{m_i}}{\sum_k U_{m_k}}. \tag{10}$$

Consider now our set of solutions, logarithmic functions in the form $U_i(E_i) = U_{m_i} \log_a(\frac{E_i}{b}) + c$, $U_{m_i} > 0$, $a > 1$, $b > 0$, $c \in R$. From Equation 1, we obtain: $\frac{U_{m_1}}{E_1^* \ln(a)b} = \frac{U_{m_2}}{E_2^* \ln(a)b} = \ldots = \frac{U_{m_n}}{E_n^* \ln(a)b}$. This way, $\frac{U_{m_1}}{E_1^*} = \frac{U_{m_2}}{E_2^*} = \ldots = \frac{U_{m_n}}{E_n^*}$. Analogously to Equation 3, it is easy to see that $\frac{U_{m_1}}{E_1^*} = \frac{U_{m_2}}{E_2^*} = \ldots = \frac{U_{m_n}}{E_n^*} = \frac{\sum_k U_{m_k}}{\sum_k E_k^*}$.

Diving all terms by the right hand side term, we obtain $\frac{E_i^*}{\sum_k E_k^*} = \frac{U_{m_i}}{\sum_k U_{m_k}}$, and so, relativistically written, $e_i^* = \frac{U_{m_i}}{\sum_k U_{m_k}}$. This effort allocation provides the same rewards as the initial option sampling in Equation 10, i.e., $e_i^* = u_i(e_i^{\text{eval}})$. This proves that contingencies that can be characterized with a set of logarithmic functions, $U(E) = U_m \log_a(\frac{E}{b}) + c$, render efficient decision making optimal in a single evaluation step.

Efficiency and diminishing marginal utility

Let us now work toward a more general conclusion. Consider decision options for which gain is characterized by a set of power functions, $U_i(E_i) = U_{m_i} E_i^{\alpha}$. From Equation 1, we obtain: $\frac{\alpha U_{m_1}}{E_1^{*\ (1-\alpha)}} = \frac{\alpha U_{m_2}}{E_2^{*\ (1-\alpha)}} = \ldots = \frac{\alpha U_{m_n}}{E_n^{*\ (1-\alpha)}}$. This can be rewritten as $\frac{U_{m_1}^{\frac{1}{1-\alpha}}}{E_1^*} = \frac{U_{m_2}^{\frac{1}{1-\alpha}}}{E_2^*} = \ldots = \frac{U_{m_n}^{\frac{1}{1-\alpha}}}{E_n^*}$. Analogously to the previous case, this equation is also equal to $\frac{U_{m_1}^{\frac{1}{1-\alpha}}}{E_1^*} = \frac{U_{m_2}^{\frac{1}{1-\alpha}}}{E_2^*} = \ldots = \frac{U_{m_n}^{\frac{1}{1-\alpha}}}{E_n^*} = \frac{\sum_k U_{m_k}^{\frac{1}{1-\alpha}}}{\sum_k E_k^*}$. Dividing the equation by the right hand side and reorganizing, we obtain $e_i = \frac{U_{m_i}^{\frac{1}{1-\alpha}}}{\sum_k U_{m_k}^{\frac{1}{1-\alpha}}}$. Thus, in this case, $e_i^* = u_i(e_i^{\text{eval}})$ when $\alpha \to 0$. Figure 3d shows that for power functions, a single evaluation can provide high performance for situations with an appreciable level diminishing marginal utility, up to about $\alpha = 0.5$. To maintain performance in situations with weakly diminishing returns, the efficient strategy must be repeated, i.e., $e_i^{\text{next}} = u_i(e_i^{\text{prev}})$ (Figure 3d). This finding provides an important insight regarding diminishing marginal utility in the derivation process leading to the optimal strategy (Figure 1). Specifically, EDM ($e_i^{\text{harvest}} = u_i(e_i^{\text{eval}})$), which is to approach the matching equilibrium of Equation 2, is subjected to even stronger requirements on diminishing marginal utility than the individual fractions of the molar equilibrium. In particular, a previous study (***Kubanek, 2017***) found that the same power functions considered here render Equation 2 optimal for any $0 < \alpha < 1$. In comparison, the molecular EDM strategy requires more strongly diminishing returns (Figure 3d) to yield adequate gain.

Let us now inspect how the level of diminishing marginal utility governs the convergence of efficient decision making in a general case. Consider the iterative form of EDM $e_i^{\text{next}} = u_i(e_i^{\text{prev}})$, or, an equivalent formulation based on Equation 3: $E_i^{\text{next}} = \frac{1}{k} U_i(E_i^{\text{prev}})$. For a gain situation that exhibits strong diminishing marginal utility, ($\frac{d^2 U_i(E_i)}{d E_i^2} \ll 0$), yet a gain that does not show negative returns ($\frac{d U_i(E_i)}{d E_i} > 0$), changes in E_i lead to only small changes in U_i. In that case, $U_i(E_i^{\text{prev}}) \approx U_i(E_i^{\text{next}})$, and $E_i^{\text{next}} \approx \frac{1}{k} U_i(E_i^{\text{next}})$. This approaches the relativistic formulation of the matching equilibrium: $E_i^* \approx \frac{1}{k} U_i(E_i^*)$, or $e_i^* \approx u_i(e_i^*)$. The above results generalize to situations in which each option can be characterized with a distinct function and situations in which the initial effort allocation is not equal but random (Figure 4b).

Simulations

Decision options

The simulations were designed to sample a large portion of the space of possible utility-effort contingencies, i.e., $U(E)$ functions. The functions were of distinct kinds (Figure 4), used distinct magnitudes, and distinct level of diminishing marginal utility. Specifically, logarithmic, hyperbolic, and power functions in the forms $U(E) = U_m \frac{\ln(E+1)}{\ln(\frac{1}{\alpha}+1)}$, $U(E) = U_m \frac{E}{E+\alpha}$, and $U(E) = U_m E^{\alpha}$, respectively,

were evaluated. The logarithmic functions used $E+1$ in their argument to avoid negative utility, and $\ln(\frac{1}{\alpha}+1)$ in the denominator to implement logarithm base values $b=\frac{1}{\alpha}+1>1$.

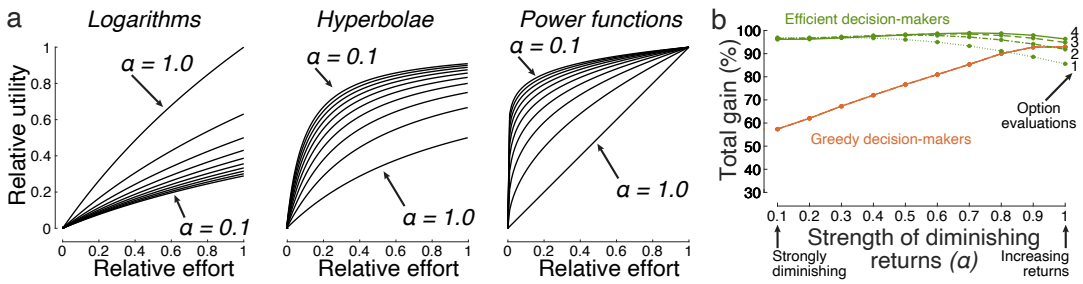

Figure 4. Efficient Decision Making and the level of diminishing marginal utility. (**a**) The utility-effort contingencies tested. Logarithmic, hyperbolic, and power functions in the forms $U(E)=U_m\frac{\ln(E+1)}{\ln(\frac{1}{\alpha}+1)}$, $U(E)=U_m\frac{E}{E+\alpha}$, and $U(E)=U_mE^{\alpha}$, respectively, were used. In these functions, the variable α quantifies the level of diminishing marginal utility, and ranged from 0.1 (strongly diminishing) to 1.0 (least-diminishing; linear for the power function), in steps of 0.1. The individual plots are scaled by these 10 values of α in the indicated order. In addition, the variable U_m in these functions scales the magnitude of each contingency, ranging from 0 to 1. (**b**) Average performance of efficient decision making and its greedy counterpart as a function of the level of diminishing marginal utility. Same format as in Figure 3d, but now for any combination of logarithmic, hyperbolic, and power functions shown in **a**. Specifically, each option (2 up to 10; 1000 evaluations each) was randomly assigned a logarithmic, hyperbolic, or power functions with the diminishing marginal utility level α indicated on the abscissa. Strong diminishing returns provide an almost immediate convergence onto the optimum. Iterative application of EDM greatly improves performance in cases with weakly diminishing or increasing returns (see number of option evaluations in black).

Choice situations

The study simulated a large number of choice situations to test the generalizability of the findings beyond those derived analytically above. The situations featured $n=2,3,\ldots 10$ options. For each option within the set of n, the specific $U(E)$ function (logarithmic, hyperbolic, or power), its magnitude U_m, and the saturation parameter α were selected randomly. U_m was selected randomly from a uniform distribution over the range from 0 to 1. The diminishing marginal utility level α was selected randomly from the set $\{0.1, 0.2, \ldots 1.0\}$. The randomization was performed 1000 times for each n, yielding a total of 9000 choice situations.

Optimizing agent

In each of the above 9000 choice situations, the optimizing agent aimed to maximize its gain within a constrained distribution of effort devoted to each option. Specifically, the agent was maximizing $\sum_i U_i(E_i)$ subject to $\sum_i E_i = E_{\max} = 1$. The problem was solved using constrained nonlinear multivariable optimization (function fmincon in Matlab, using the default interior-point algorithm). Step and solution tolerances were set to 10^{-4} to decrease the number of necessary option evaluations and so accelerate the computations. The algorithm converged onto a solution, within the tolerances, in a finite number of options evaluations in each decision case (see Figure 2 for the number of option evaluations necessary to reach a specific performance level). This is expected since the utility-effort functions were well behaved and noise-free. Notably, for noisy

forms of $U(E)$, the optimizing agent is at further disadvantage compared with EDM as it can get stuck on local optima (data not shown). The EDM using fractional representations (Equation 3), is relatively robust to noise.

Efficient Decision Making

Like the optimizing agent, the biologically feasible decision-makers distributed their effort E_i across the choice options subject to finite effort, $\sum_i E_i = E_{\max} = 1$. Their total gain $\sum_i U_i(E_i)$, is reported as a percentage of the maximum possible gain harvested by the optimizing agent. Unlike the agent, however, the decision-makers used a single or just a small number of evaluations of the decision options. Specifically, they first sampled the worth $u_i(e_i^{\text{eval}})$ of their options i by allocating initial effort e_i^{eval} to each option. Subsequently, they applied the efficient decision making strategy i.e., $e_i^{\text{harvest}} = u_i(e_i^{\text{eval}})$. Thus, the decision-makers simply allocated their effort in proportion to an initially established worth of their options. In Figure 3b, Figure 3d, and Figure 4b, this strategy was applied iteratively, i.e., $e_i^{\text{next}} = u_i(e_i^{\text{prev}})$, up to 4 times. Moreover, the analysis of Figure 3b also tested whether the strategy is sensitive to the kind of the initial effort distribution—equal ($e_1^{\text{eval}} = e_2^{\text{eval}} = \ldots = e_n^{\text{eval}}$) or entirely random. Given that an equal allocation of effort across options may not be realistic in natural or human settings, the results in Figure 3c, Figure 3d, and Figure 4b used the random allocation of initial effort. Figure 3a evaluated how the efficient strategy compares to a more general strategy, for which, for the most desirable option, $e_i^{\text{harvest}} = g u_i(e_i^{\text{eval}})$. Here, the parameter g ranged from 0.2 to 4.0. This parameter tested varying levels of greed ($g > 1$) and indifference ($g < 1$).

Acknowledgements

I thank Drs. Alex Kacelnik, Alasdair Houston, Leo Sugrue, and Julian Brown for comments on the manuscript. This work was supported by the NIH grant R00NS100986.

Code availability

The simulation code is available and documented at onetarget.us/download/EDM.

References

Baum WM. On two types of deviation from the matching law: bias and undermatching. Journal of the experimental analysis of behavior. 1974; 22(1):231–242.

Becker G. The Economic Approach to Human Behavior. University of Chicago Press; 1978.

Białaszek W, Ostaszewski P, Green L, Myerson J. On four types of devaluation of outcomes due to their costs: Delay, probability, effort, and social discounting. The Psychological Record. 2019; 69(3):415–424.

Binmore K. Rational Choice Theory: Sufficient but not necessary. American Psychologist. 1991; 46(7):797–799.

Botvinick MM, Huffstetler S, McGuire JT. Effort discounting in human nucleus accumbens. Cognitive, affective, & behavioral neuroscience. 2009; 9(1):16–27.

Buchanan JM. Cost and choice: An inquiry in economic theory. University of Chicago Press; 1978.

Charnov EL. Optimal foraging, the marginal value theorem. Theoretical population biology. 1976; 9(2):129–136.

Cobb CW, Douglas PH. A theory of production. The American Economic Review. 1928; 18(1):139–165.

Corrado GS, Sugrue LP, Seung HS, Newsome WT. Linear-Nonlinear-Poisson models of primate choice dynamics. Journal of the experimental analysis of behavior. 2005 Nov; 84(3):581–617. http://www.ncbi.nlm.nih.gov/pubmed/16596981?dopt=abstract.

Friedman M. Essays in positive economics. University of Chicago Press; 1953.

Friedman M, Savage LJ. The utility analysis of choices involving risk. Journal of political Economy. 1948; 56(4):279–304.

Gossen HH. Entwickelung der Gesetze des menschlichen Verkehrs, und der daraus fließenden Regeln für menschliches Handeln. Friedrich Vieweg und Sohn; 1854.

Hamblin S, Giraldeau LA. Finding the evolutionarily stable learning rule for frequency-dependent foraging. Animal Behaviour. 2009; 78(6):1343–1350.

Harley CB. Learning the evolutionarily stable strategy. Journal of theoretical biology. 1981; 89(4):611–633.

Herrnstein RJ. Relative and absolute strength of response as a function of frequency of reinforcement. Journal of the experimental analysis of behavior. 1961 Jul; 4:267–72. http://www.ncbi.nlm.nih.gov/sites/entrez?Db=pubmed&Cmd=Retrieve&list_uids=13713775&dopt=abstractplus, doi: 10.1901/jeab.1961.4-267.

Herrnstein RJ. The matching law: Papers in psychology and economics. Harvard University Press; 2000.

Hollis M, Nell EJ. Rational economic man: a philosophical critique of neo-classical economics. Cambridge University Press; 1975.

Houston AI, Sumida BH. Learning rules, matching and frequency dependence. Journal of Theoretical Biology. 1987; 126(3):289–308.

Houston AI, Trimmer PC, McNamara JM. Matching behaviours and rewards. Trends in Cognitive Sciences. 2021; 25(5):403–415.

Hutchinson JM, Gigerenzer G. Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet. Behavioural processes. 2005; 69(2):97–124.

Iigaya K, Ahmadian Y, Sugrue LP, Corrado GS, Loewenstein Y, Newsome WT, Fusi S. Deviation from the matching law reflects an optimal strategy involving learning over multiple timescales. Nature communications. 2019; 10(1):1–14.

Kacelnik A, Krebs JR, Bernstein C. The ideal free distribution and predator-prey populations. Trends in Ecology & Evolution. 1992; 7(2):50–55.

Karatzas I, Wang H. Utility maximization with discretionary stopping. SIAM Journal on Control and Optimization. 2000; 39(1):306–329.

Kubanek J. Optimal decision making and matching are tied through diminishing returns. Proceedings of the National Academy of Sciences. 2017; 114(32):8499–8504.

Louie K, Glimcher PW. Separating value from choice: delay discounting activity in the lateral intraparietal area. Journal of Neuroscience. 2010; 30(16):5498–5507.

McNamara J, Houston A. The application of statistical decision theory to animal behaviour. Journal of Theoretical Biology. 1980; 85(4):673–690.

Myerson J, Green L. Discounting of delayed rewards: Models of individual choice. Journal of the experimental analysis of behavior. 1995; 64(3):263–276.

Rachlin H, Battalio R, Kagel J, Green L. Maximization theory in behavioral psychology. Behavioral and Brain Sciences. 1981; 4(3):371–388.

Robinson J. Economic Philosophy. Penguin Books; 1962.

Samuelson PA. Foundations of economic analysis. Science and Society. 1948; 13(1).

Simon HA. A behavioral model of rational choice. The quarterly journal of economics. 1955; 69(1):99–118.

Simon HA. Rational choice and the structure of the environment. Psychological review. 1956; 63(2):129.

Stephens DW, Krebs JR. Foraging theory. Princeton University Press; 1986.

Stephens DW, Krebs JR. Foraging theory. Princeton university press; 2019.

Sugrue LP, Corrado GS, Newsome WT. Matching behavior and the representation of value in the parietal cortex. Science. 2004 Jun; 304(5678):1782–7. http://www.sciencemag.org/cgi/content/full/304/5678/1782, doi: 10.1126/science.1094765.

Sugrue LP, Corrado GS, Newsome WT. Choosing the greater of two goods: neural currencies for valuation and decision making. Nat Rev Neurosci. 2005 May; 6(5):363–75. http://www.nature.com/nrn/journal/v6/n5/abs/nrn1666.html;jsessionid=tk8ofxxilcb3/index.html, doi: 10.1038/nrn1666.

Sutton RS, Barto AG. Introduction to reinforcement learning. MIT Press; 1998.

Tracy ND, Seaman Jr JW. Properties of evolutionarily stable learning rules. Journal of theoretical biology. 1995; 177(2):193–198.

Vaughan Jr W, Herrnstein RJ. Stability, melioration, and natural selection. Advances in behavioral economics. 1987; 1:185–215.

Zipf GK. Human behavior and the principle of least effort: an introduction to human ecology. Addison-Wesley Press; 1949.