Optimal decision making and matching are tied through diminishing returns

Jan Kubanek[a,1]

[a]Department of Neurobiology, Stanford University School of Medicine, Stanford, CA 94305

Edited by Ranulfo Romo, Universidad Nacional Autonóma de México, Mexico City, D.F., Mexico, and approved July 3, 2017 (received for review March 1, 2017)

How individuals make decisions has been a matter of long-standing debate among economists and researchers in the life sciences. In economics, subjects are viewed as optimal decision makers who maximize their overall reward income. This framework has been widely influential, but requires a complete knowledge of the reward contingencies associated with a given choice situation. Psychologists and ecologists have observed that individuals tend to use a simpler "matching" strategy, distributing their behavior in proportion to relative rewards associated with their options. This article demonstrates that the two dominant frameworks of choice behavior are linked through the law of diminishing returns. The relatively simple matching can in fact provide maximal reward when the rewards associated with decision makers' options saturate with the invested effort. Such saturating relationships between reward and effort are hallmarks of the law of diminishing returns. Given the prevalence of diminishing returns in nature and social settings, this finding can explain why humans and animals so commonly behave according to the matching law. The article underscores the importance of the law of diminishing returns in choice behavior.

choice behavior | rational choice theory | economic maximization | neuroeconomics | matching law

People's decisions define the future of individuals and social groups. How to suitably model, understand, and predict individuals' choice behavior has therefore been a matter of intense research efforts involving multiple disciplines.

Economic models provide normative prescriptions of how individuals should make decisions. According to these models, individuals make decisions in order to maximize their expected reward, utility, or income (1–3). For example, in prospect theory (4), subjects maximize the expected utility of potential decision outcomes. The expected utility is computed as the sum of the outcomes' values weighted by the probabilities that the individual outcomes will occur. Within such maximization models, Bayesian decision theories can be used to dictate how subjects should optimally compute the individual probability terms (5). Despite the normative appeal of such maximization models, it has been unclear whether organisms are capable of implementing and acting on the complex computations prescribed by these models (6–8).

Researchers in psychology, ecology, sociology, and neuroscience found evidence for a relatively simpler model for decision making. It has been found that decision makers tend to distribute their behavior in proportion to relative rewards associated with their options (3, 8–16). The match between the behavioral and reward distributions, $\frac{B_i}{B_1+B_2+\ldots+B_n} = \frac{R_i}{R_1+R_2+\ldots+R_n}$, where B_i is the rate of behavior allocated at option i, and R_i is the corresponding rate of the obtained reward, has come to be known as the "matching law" (8, 9, 11, 13, 14). Analogously to economic models, R_i can also represent utilities. This way, matching may capture a wide range of decision conditions while remaining relatively compact (14, 17). However, matching has been criticized for lacking a theoretical basis; matching is an empirical phenomenon (17–19).

Which of these models is the more appropriate to capture and predict choice behavior has been a subject of substantial debate (7, 8, 18–22). In some cases, subjects maximize and do not match (23, 24), whereas in other cases subjects match even though maximization would be a better strategy (3, 20, 25, 26). Psychologists have compared matching and maximization in tasks that used specific schedules of reinforcement (24, 27, 28), but whether the two models can be linked analytically at a general level has remained elusive. The present article turns matching and maximization face to face and at a general level. By doing so, it identifies a connection between the economic and psychological frameworks, and the nature of the connection provides an explanation for why humans and animals so often follow the matching strategy.

PSYCHOLOGICAL AND COGNITIVE SCIENCES

Results

Optimal Decision Making. An optimal decision maker distributes her effort across options such as to maximize the total harvested reward. When effort E_i allocated at option i yields reward rate $R_i(E_i)$, the total reward rate the decision maker obtains for a specific distribution of effort among her n options amounts to $R_1(E_1) + R_2(E_2) + \ldots + R_n(E_n)$. The total effort that an individual can invest is limited, $\sum_i E_i = E_{max}$. Given that, reward optimum is attained (*Methods*) when $\frac{dR_1(E_1)}{dE_1} = \frac{dR_2(E_2)}{dE_2} = \ldots = \frac{dR_n(E_n)}{dE_n}$. A decision maker can maximize her total reward by equalizing marginal reward per effort $\frac{dR(E)}{dE}$ across her options. R, the expected rate of reward, is in this article for simplicity referred to as "reward."

Matching Behavior. Humans and animals often follow a matching strategy (3, 8–16), distributing their behavior B_i in proportion

Significance

Decisions critically affect the well-being of individuals and societies. However, how to suitably model, understand, and predict people's decisions has been a long-standing challenge. Economic models view individuals as optimal decision makers who maximize their overall reward income. Psychologists and ecologists have observed that decision makers tend to use a relatively simpler "matching" strategy, distributing their behavior in proportion to relative worth of their options. This article demonstrates that matching can be an optimal strategy for decision makers when the rewards associated with their options diminish with the invested effort, a relationship known as the law of diminishing returns. Because diminishing returns are prevalent in nature and social settings, the commonly observed matching behavior is not only simple but also efficient and rational.

Author contributions: J.K. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

The author declares no conflict of interest.

This article is a PNAS Direct Submission.

[1]Email: kubanek@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1703440114/-/DCSupplemental.

to relative rewards R_i associated with their options, $\frac{B_i}{B_1+B_2+\ldots+B_n} = \frac{R_i}{R_1+R_2+\ldots+R_n}$. This "matching" equation (9, 11) can be equivalently written as $\frac{R_1}{B_1} = \frac{R_2}{B_2} = \ldots = \frac{R_n}{B_n}$. Furthermore, noting that behavior B embodies effort E, matching can be stated as $\frac{R_1}{E_1} = \frac{R_2}{E_2} = \ldots = \frac{R_n}{E_n}$. It becomes apparent that when subjects match, they equalize average reward per effort $\frac{R}{E}$ across their options. When the value of an option i, V_i, is defined as the ratio of the obtained reward per the invested effort, $V_i = \frac{R_i}{E_i}$, matching equalizes the values $V_i = V$ across the choice options.

Maximization–Matching Relation. Given these observations, maximization and matching align when marginal reward per effort associated with an option is a strictly monotonic function g of average reward per effort associated with the option, $\frac{dR}{dE} = g\left(\frac{R}{E}\right)$ (*Methods*). The marginal and average quantities for an example relationship between reward and effort are illustrated as dotted and dashed lines in Fig. 1, respectively.

This formulation empowers us to identify the relationships (henceforth referred to as contingencies) between reward and effort $R(E)$ for which the two approaches provide an equal amount of reward. The contingencies emerge simply by defining a particular form of the strictly monotonic function g. There are only two requirements (*Methods*) on g to yield contingencies for which matching delivers maximal reward: (*i*) $g(V)$ must have its derivative greater than zero and (*ii*) it must be $g(V) < V$. Simple forms of g (*Methods*) yield contingencies including $R(E) = aR\ln(E) + c$ with $a > 0$, $R(E) = RE^a$ with $0 < a < 1$, and $R(E) = \frac{E}{a+cE}$ with $a, c > 0$, where R is the reward associated with an option, E is the invested effort, and a and c are constants.

These derived contingencies provide tight fits to the relationship between reward and effort in a dominant task used to study choice behavior (8, 9, 13). In particular, in the variable-interval schedule task (Fig. 2), an increase in a subject's effort generally leads to an increase in the harvested reward. However, at a certain point the reward saturates and shows diminishing returns with additional effort (Fig. 2). The derived $R(E)$ contingencies for which matching is optimal provide excellent fits to the diminishing-returns character of this task, explaining 97.8%, 95.9%, and 99.7% of variance, respectively (Fig. 2). It is apparent that these solutions, just like the reward–effort profile of the variable-interval schedule task, exhibit diminishing returns.

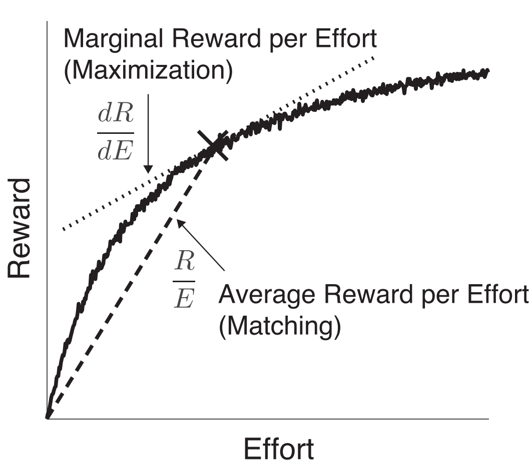

Fig. 1. Relationship between reward maximization and matching. Reward-maximizing behavior and matching align when marginal ($\frac{dR}{dE}$) and average ($\frac{R}{E}$) reward per effort for each option are related through a strictly monotonic function. This function can be distinct from identity, and so the marginal (dotted line) and average (dashed line) quantities can have different slopes.

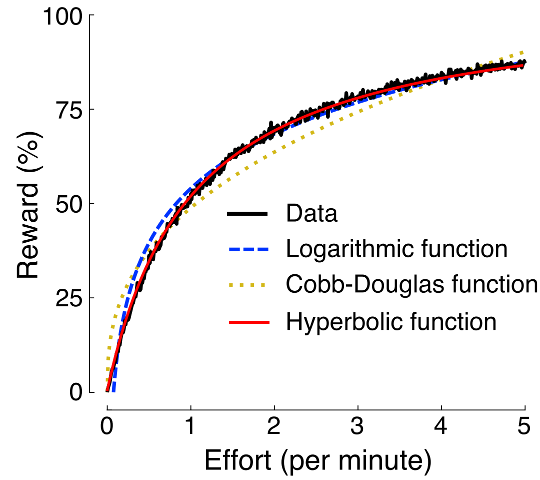

Fig. 2. Relationships between reward and effort for which matching is optimal exhibit diminishing returns. Choice behavior has often been studied using the variable-interval schedule task. In this task, a single reward is scheduled at an option at a specific rate, such as, on average, one per minute here. Once scheduled, the reward is available until a subject chooses the option. A simulation of the reward harvested in this task (black curve) shows that when a subject chooses an option at the same rate as the schedule rate (once per minute), reward is delivered on average in 50% of the cases, as expected. Furthermore, as expected, investing more effort leads to more reward. Nonetheless, at a certain point, there are diminishing returns. The diminishing-returns character of the data can be well captured with the derived reward–effort contingencies, exemplified by three particular functions (colored curves). A logarithmic function ($R(E) = aR\ln(E) + c$) explains 97.8% of variance in the data, a Cobb–Douglas function ($R(E) = RE^a$) 95.9% of variance, and a hyperbola ($R(E) = \frac{E}{a+cE}$) 99.7% of variance. Similarly tight fits are observed also for other reward scheduling rates.

This, in fact, turns out to be a general finding. For matching to provide a reward maximum, the reward–effort contingencies of the choice options must exhibit diminishing returns (see *Methods* for proof). This finding ties the matching law, deemed by many a fundamental principle underlying choice behavior (3, 11, 12, 14) with economic maximization through a surprising link—the law of diminishing returns.

This conclusion holds also in environments in which rewards are stochastic and for maximization metrics that explicitly incorporate probabilistic outcomes. For instance, when rewards R_i are delivered with probabilities p_i, an optimal decision maker distributes her effort to maximize the expected value of reward over the individual options, $E[R] = p_1R_1(E_1) + p_2R_2(E_2) + \ldots + p_nR_n(E_n)$. Diminishing returns turn out to be a critical requirement for matching to be optimal also under this optimization criterion (*SI Methods*). In addition, the $E[R]$ criterion can be rewritten to accommodate expected utility (4, 5), and the same connection between matching and diminishing returns is observed (*SI Methods*).

Discussion and Conclusions

The matching law has been an influential description of choice behavior of animals and humans (3, 8–15), but how this principle relates to the maximization apparatus of economic theories and why individuals so often adopt matching have been unclear. This article notes that reward depends on the invested effort and identifies the reward–effort contingencies for which matching is an optimal choice strategy. It is found that the contingencies for which matching provides a reward maximum share one striking commonality: They exhibit diminishing returns. In these cases (e.g., Fig. 2), the harvested reward increases as a function of the invested effort up to a point where additional investment confers a relatively small marginal benefit. The law of diminishing returns (29, 30) turns out to be an intriguing link that connects the psychological matching law with economic maximization.

The finding that matching is crucially based on diminishing returns was not obviously derivable from previous literature. The present study arrives at this finding by developing a framework that relates matching and optimization at a general level, by providing a complete space of solutions for which matching is optimal and by deriving the requirements for matching to attain reward maxima. Previous studies identified only two specific solutions, $R(E) = RE^a$ and $R(E) = \frac{E}{a+cE}$ (20, 21, 27). In addition, these specific solutions were identified only as those for which matching can attain a critical point in the reward landscape, which can be a maximum but also a minimum or a saddle point. The present study shows that framing the problem at a general level and providing the requirements for matching to attain reward maxima prove crucial for uncovering the involvement of diminishing returns (proof in *Methods*).

The finding that matching critically rests on situations with diminishing returns brings us to the question of how commonly diminishing returns figure in nature and human society. The law of diminishing returns (29, 30) traces its history back to Turgot who discovered that agricultural output progressively decreases with increasing quantities of invested capital and labor (31). The idea has subsequently been elaborated by economists such as Malthus and Ricardo. The law of diminishing returns now lies at the heart of many branches in economics, including production, investment, and economic growth theories (32–35). For example, the present article identifies the Cobb–Douglas function ($R(E) = RE^a$), often used to capture the diminishing returns on input labor in regard to production output, as one of the solutions for which matching is optimal. For matching to be optimal, the exponent a of effort (labor in this case) must be $0 < a < 1$ (*Methods*), just as prescribed by the Cobb–Douglas function. This constitutes diminishing returns. Exemplified in production, if two or more production processes can be described by a Cobb–Douglas function with equal exponents a, then matching guarantees an optimal distribution of labor between the processes. In this case, to maximize the total production, the matching strategy allows a manager to simply equalize the average production per total labor allocated within each production process.

The findings of this article also apply to ecology. Consider a predator who must decide how to distribute her foraging effort between sources of prey. Foraging or harvest situations commonly involve diminishing returns (36). For example, either the predator gradually depletes the prey or the regeneration of a resource saturates due to factors such as crowding. This article shows that situations of diminishing returns allow the predator to distribute her effort effectively according to the matching law, equalizing the value (the obtained reward per invested effort) of each source. If in source 2 she obtains twice the amount of prey per foraging effort as in source 1, she spends twice the amount of effort on source 2 compared with source 1. The major benefit of this approach is that it is simple. The predator visits the sources to determine their value and distributes her effort to harvest equal value from both (37, 38). This is in contrast to a reward-maximizing agent that must learn about the outcomes of all possible allocations of effort across the sources and compute derivatives to evaluate the marginal reward per effort across the sources (39). In situations of diminishing returns like this, this article shows that matching represents an effective heuristic to maximizing reward income.

Diminishing returns also apply to situations that rest on temporal, financial, and mental effort. For instance, when effort involves time ($E = T$), the value terms take the form $V = \frac{R}{T}$, which represents hyperbolic temporal discounting (40). In this regard, this article reveals that in situations in which the invested time reaches diminishing returns, the hyperbolic discounting of temporal effort embodies an optimal value function for decisions concerned with how to allocate time across choice options. Analogous reasoning applies to monetary investment and mental effort. In situations of diminishing returns, the present article suggests that making a choice according to the matching law constitutes a good strategy.

Matching is not a ubiquitous phenomenon. There are situations, such as those modeled by random (variable) ratio schedules, in which the reward rate $R(E)$ increases proportionally with effort E. In such situations with nondiminishing returns, according to this article, subjects should not match (*Methods*). Indeed, subjects in such cases commonly converge on almost exclusively choosing the richer alternative (23, 24). Therefore, the presence or absence of diminishing returns in a given task can be used as an indicator of whether subjects should or should not exhibit matching.

Diminishing returns embody a necessary condition for matching to provide maximal total reward harvested. This is not at the same time a sufficient condition; one can find examples of saturating functions for which matching does not imply a reward maximum. Such saturating functions can nonetheless be approximated with functions derived using specific forms of the generator function g. Thus, the extent to which diminishing returns present a sufficient condition for matching to maximize reward rests on the extent and variety of the functions that can be generated using g.

In addition to illuminating the relationship between matching and reward maximization, this article also contributes to the body of research on effort-based decision making, in two ways. First, in the present framework, effort is treated as a resource instead of a variable with negative valence. Second, it is shown that computing value V of an option as reward per effort, $V = \frac{R}{E}$, and operating on such value representation through matching can approach, for contingencies of diminishing returns, maximal reward. The fractional representation $V = \frac{R}{E}$ circumvents the difficulty to express reward R and effort E on the same scale (41). Intriguingly, the fractional representation of value, $V = \frac{R}{E}$, has been found to be encoded in specific regions of the brain (42, 43).

Reward maximization represents a behavioral equilibrium characterized by equalized marginal reward per unit of effort across the choice options. Which optimization strategy may result in such an equilibrium has been a matter of debate (16, 18, 44). In stochastic environments, one of the main candidate frameworks that can lead to reward maximization has been the Bayesian decision theory. According to this formalism, subjects make decisions such as to maximize the expected utility, which incorporates probability terms that model uncertain relationships between decisions and outcomes (5, 45). The Bayesian framework provides an optimal prescription for how individuals should update their probability estimates given prior experience and recent evidence. Although the neuronal computations underlying Bayesian inference appear to be biologically plausible (5, 46), it remains to be seen whether such computations can be combined with utility representations to provide the maximization metrics necessary to guide optimal decisions in complex choice situations (4, 5).

Matching also represents a behavioral equilibrium, characterized by equalized average reward per effort across the choice options. Several behavioral strategies have been shown to result in matching (16, 47–50). One of the leading candidates, directly derived from matching, has been melioration (47, 51). According to melioration, decision makers assess, over a certain time period, the value (reward per effort) of each option and adjust their effort, with certain frequency, to the option with the highest value. Equilibrium is achieved and subjects match when the values (reward per effort) are equalized across the options. In contrast, optimization, the process that leads to reward maximization, continuously reallocates effort to the option with the highest marginal reward per effort. Equilibrium is achieved when the marginals are equal across the choice options (*Optimal Decision Making*). There are two major advantages of melioration over optimization. First, evaluating reward that has cumulated over a certain time period reduces noise in the reward income in stochastic environments. Second, because melioration operates on a certain time period, effort can be adjusted with a frequency corresponding to that time period; during optimization, effort is adjusted at each point in time. Evaluating cumulative values every so often is a biologically much more plausible strategy compared with computing local derivatives at each point in time. This paper shows that making decisions based on a certain time period—inherent to melioration and matching—can constitute a good strategy in choice environments with diminishing returns.

Deciding between different classes of options, such as whether to cook at home or eat out at a restaurant, involves a comparison of utilities and efforts associated with the options. In this regard, melioration as a simple behavioral strategy can be generalized to represent utilities instead of reward rates (14, 17). In this generalized model, subjects choose the option that provides the highest utility per effort (or per economic cost). In equilibrium, such a strategy leads to matching of effort (or financial resources) to the relative utilities.

In sum, the present article links two dominant frameworks of choice behavior and finds that matching can be an efficient and effective instantiation of economic maximization as long as the choice environment features diminishing returns. The observations that humans and animals so often behave according to the matching law now find footing in the law of diminishing returns. In light of diminishing returns, matching becomes an efficient heuristic to optimal decision making.

Methods

Reward and Effort in Variable-Interval Schedules of Reinforcement. The relationship between reward and effort in Fig. 2 was obtained using a simulation of the variable-interval schedule task. In this task, a reward of a certain magnitude is delivered at random intervals but at a constant overall rate. For example, the data shown in Fig. 2 use a rate of 1 reward per minute. The decision whether to schedule a reward or not at a certain time is governed by a Poisson process; each point in time has an equal probability that a reward will be scheduled. Once a reward is scheduled, it remains available until a subject (a computer in this case) harvests it. The abscissa of the plot (i.e., effort) provides the number of times per minute that the subject checked whether there was a reward. The subject's decision to check is also driven by a Poisson process in which the probability of checking is the same at any point in time.

Optimal Decision Making. This section derives how effort should be distributed among choice options to maximize the total reward harvested, $R_1(E_1) + R_2(E_2) + \ldots + R_n(E_n)$, subject to the constraint that total effort of a decision maker is limited, $\sum_i E_i = E_{\max}$. The criterion constitutes the total reward to be maximized, independently of effort. Effort serves as a means to maximize the total reward. This problem can be solved using the method of Lagrange multipliers. The Lagrangian is in this case formulated as

$$\mathcal{L}(E_1, E_2, \ldots, E_n, \lambda) = R_1(E_1) + R_2(E_2) + \ldots + R_n(E_n) + \lambda(E_{\max} - \sum_i E_i). \quad [1]$$

Setting the partial derivatives to 0, we get

$$\frac{\partial \mathcal{L}}{\partial E_i} = \frac{dR_i(E_i)}{dE_i} - \lambda = 0, \quad [2]$$

and so

$$\frac{dR_i(E_i)}{dE_i} = \lambda$$

for a certain λ. This means that

$$\frac{dR_1(E_1)}{dE_1} = \frac{dR_2(E_2)}{dE_2} = \ldots = \frac{dR_n(E_n)}{dE_n}. \quad [3]$$

Eq. **3** dictates how effort should be allocated to attain a critical point in the reward landscape. This can be a maximum, a minimum, or a saddle point. To obtain a reward maximum, the leading principal minors of the bordered Hessian matrix corresponding to Eq. **1**, evaluated at critical points, must alternate in sign, with the first minor (of order 3) showing a positive sign (34).

Let us label $\frac{d^2 R_i(E_i)}{dE_i^2} = R_i''(E_i)$ at a certain critical point E_i. The bordered Hessian for Eq. **1** is

$$H^B(E_1, E_2, \ldots, E_n) = \begin{bmatrix} 0 & -1 & -1 & \ldots & -1 \\ -1 & R_1''(E_1) & 0 & \ldots & 0 \\ -1 & 0 & R_2''(E_2) & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & 0 & R_n''(E_n) \end{bmatrix}.$$

For two options, the leading principal minor of order 3 is equal to

$$\det H^B(E_1, E_2) = -(R_1''(E_1) + R_2''(E_2)).$$

This leading principal minor must be positive, so it must be

$$R_1''(E_1) + R_2''(E_2) < 0.$$

This result leads to two important arguments that are used below. First, for a critical point to embody a reward maximum—so that the above equation holds—at least for one option and at least for some value of E_i the second derivative $R_i''(E_i)$ must be $R_i''(E_i) < 0$. Second, if for any E_i and for both options $R_i''(E_i) < 0$, then the above equation always holds and the critical point is guaranteed to be a maximum.

The same two arguments also hold for choice situations that feature three and more options. For example, for three options, the leading principal minor of order 4 is equal to

$$\det H^B(E_1, E_2, E_3) = -(R_1''(E_1)R_2''(E_2) + R_1''(E_1)R_3''(E_3) + R_2''(E_2)R_3''(E_3))$$

and it must be negative. Thus, it must be

$$R_1''(E_1)R_2''(E_2) + R_1''(E_1)R_3''(E_3) + R_2''(E_2)R_3''(E_3) > 0,$$

and the conditions associated with the principal minors of order 3 must also hold. Thus, again, for a critical point to represent a reward maximum, at least for one option and at least for some value of E_i must $R_i''(E_i) < 0$. And, if for any E_i and for all three options $R_i''(E_i) < 0$, then the above equation evaluates positive, and so the critical point is a maximum. Analogously, for four options, it must be

$$R_1''(E_1)R_2''(E_2)R_3''(E_3) + R_1''(E_1)R_3''(E_3)R_4''(E_4) + R_1''(E_1)R_3''(E_3)R_4''(E_4) + R_2''(E_2)R_3''(E_3)R_4''(E_4) < 0,$$

and the conditions associated with the principal minors of lower orders must also hold. Here, the sum of all triplets must be negative; for five options, the sum of all quadruplets must be positive, and so on. Each additional option adds an additional term in the multiples; if each such term is negative (i.e., the second derivative $R_i''(E_i) < 0$ for all i), the expression flips sign, as required. Thus, it is apparent that the two arguments we initially made for two options hold for any number of options in this constraint-optimization problem.

Matching Behavior. Behavior according to the matching law, $\frac{B_i}{B_1 + B_2 + \ldots + B_n} = \frac{R_i}{R_1 + R_2 + \ldots + R_n}$, where B_i is the rate of behavior allocated at option i and R_i is the corresponding rate of the obtained reward, can be equivalently written as $\frac{R_1}{B_1} = \frac{R_2}{B_2} = \ldots = \frac{R_n}{B_n}$. Furthermore, noting that behavior B embodies effort E, the matching law can be stated as $\frac{R_1}{E_1} = \frac{R_2}{E_2} = \ldots = \frac{R_n}{E_n}$.

Maximization–Matching Relation. Matching and optimization align when marginal reward $\frac{dR}{dE}$ associated with an option is a strictly monotonic function g of the corresponding matching term $\frac{R}{E}$:

$$\frac{dR}{dE} = g\left(\frac{R}{E}\right). \quad [4]$$

To see why, incorporating Eq. **4** into Eq. **3** leads to $\frac{dR_1(E_1)}{dE_1} = g\left(\frac{R_1(E_1)}{E_1}\right) = \frac{dR_2(E_2)}{dE_2} = g\left(\frac{R_2(E_2)}{E_2}\right) = \ldots = \frac{dR_n(E_n)}{dE_n} = g\left(\frac{R_n(E_n)}{E_n}\right)$. A strictly monotonic function is invertible. Inverting g, we obtain

$$\frac{R_1(E_1)}{E_1} = \frac{R_2(E_2)}{E_2} = \ldots = \frac{R_n(E_n)}{E_n},$$

which represents the matching law.

Having turned the matching law and optimization face to face, we can identify the contingencies between reward and effort, $R(E)$, for which these two models align.

Reward–Effort Contingencies for Estimated Reward Returns. Let us assign the matching terms $\frac{R_i}{E_i}$ a value function $V = \frac{R_i}{E_i}$. In this expression, the reward R_i associated with an option can be either an estimate or an actual return. Let us first deal with the former case and label a subject's estimate of R_i as $\hat{R}_i$. For a simple case of $g(V) = aV$, Eq. **4**, now becoming $\frac{dR_i(E_i)}{dE_i} = a\frac{\hat{R}_i}{E_i}$, yields the following reward–effort contingency:

$$R_i(E_i) = a\hat{R}_i \ln(E_i) + c_i. \quad [5]$$

Here, $\hat{R}_i$ is the subject's estimate of the reward associated with option i, and c_i and a are arbitrary constants.

Reward–Effort Contingencies for Actual Reward Returns. Let us now consider the cases in which subjects match their response distribution to the actual returns obtained from choosing an option. In such cases, the value function takes the form $V(R_i(E_i), E_i) = \frac{R_i(E_i)}{E_i}$, where $R_i(E_i)$ is the actual reward obtained for effort E_i exerted at option i. For $g(V) = aV$, Eq. **4**, now becoming $\frac{dR_i(E_i)}{dE_i} = a\frac{R_i(E_i)}{E_i}$, yields the solution $R_i(E_i) = c_iE_i^a$. For c_i interpreted as the amount of reward R_i obtained for unitary effort $E_i = 1$, the solution becomes

$$R_i(E_i) = R_iE_i^a. \quad [6]$$

For this function, matching yields a reward maximum when the exponent a is $0 < a < 1$ (see below). Thus, Eq. **6** embodies the Cobb–Douglas function that was originally applied to model diminishing returns in production output as a function of the input labor (32). This function has also been identified to relate reinforcement and behavior in psychology (21).

As another example, for $g(V) = aV^2$, Eq. **4** generates a reward–effort contingency

$$R_i(E_i) = \frac{E_i}{a + c_iE_i}. \quad [7]$$

This solution corresponds to a derivation of a feedback function for the variable-interval schedule task (52).

All other reward–effort contingencies can be generated by defining a specific $g(V)$. The admissible forms of this function are provided next.

Reward Maxima and Diminishing Returns. Matching provides a critical point in the reward landscape for every solution of Eq. **4**. It will be shown here that whether the critical point is a maximum, a minimum, or a saddle point depends solely on the properties of the generator function $g(V)$. Furthermore, it will be shown that for matching to deliver a reward maximum, the reward–effort profiles of all choice options must show diminishing returns. The proof follows from evaluating the second derivative of $R(E)$.

Estimated reward returns. Applying the chain rule, the derivative of Eq. **4** with respect to effort for estimated rewards $\hat{R}$ is equal to

$$\frac{d^2R(E)}{dE^2} = g'\left(\frac{\hat{R}}{E}\right)\left(-\frac{\hat{R}}{E^2}\right).$$

Optimal Decision Making showed that for our constraint optimization problem to yield a reward maximum, it must be $\frac{d^2R(E)}{dE^2} < 0$ for at least one option and for at least some value of E. To meet that requirement, because in the above expression $-\frac{\hat{R}}{E^2} < 0$, it must be $g'(\frac{\hat{R}}{E}) > 0$ for at least one option and for at least some E. But because of matching, the g' argument $\frac{\hat{R}}{E}$ is equal for all options because $\frac{\hat{R}_1}{E_1} = \frac{\hat{R}_2}{E_2} = \ldots = \frac{\hat{R}_n}{E_n} = V$. It follows that it must be $g'(V) > 0$, and this value is equal for all options. Thus, whether matching yields a reward maximum or a minimum depends solely on this requirement on g. Furthermore, when this requirement is fulfilled, it is apparent that $\frac{d^2R_i(E_i)}{dE_i^2} < 0$ for all options and for any values of E_i. This proves that the reward–effort contingencies for all options are negatively accelerated, i.e., show diminishing returns, and they do so for all values of effort.

Actual reward returns. The proof proceeds in a similar fashion for actually obtained reward returns $R(E)$. In this case, the derivative of Eq. **4** with respect to effort is

$$\frac{d^2R(E)}{dE^2} = g'\left(\frac{R(E)}{E}\right)\left(\frac{\frac{dR(E)}{dE}E - R(E)}{E^2}\right)$$
$$= g'\left(\frac{R(E)}{E}\right)\left(\frac{\frac{dR(E)}{dE} - \frac{R(E)}{E}}{E}\right).$$

Here, according to Eq. **4**, $\frac{dR(E)}{dE} = g(\frac{R(E)}{E})$. Thus,

$$\frac{d^2R(E)}{dE^2} = g'\left(\frac{R(E)}{E}\right)\left(\frac{g(\frac{R(E)}{E}) - \frac{R(E)}{E}}{E}\right).$$

The beauty of matching shines through this equation. Given that $E > 0$, it is apparent that the sign of the right side of the equation is solely a function of the matching term $V = \frac{R(E)}{E}$:

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

Again, *Optimal Decision Making* showed that for a critical point to represent a reward maximum, it must be $\frac{d^2R(E)}{dE^2} < 0$ for at least one option and at least for some value of E. To meet that requirement, and taking into account $E > 0$, there are two possibilities: (*i*) $g'(V) > 0$ and $g(V) < V$ and (*ii*) $g'(V) < 0$ and $g(V) > V$. The second possibility is inadmissible because for a general $V > 0$, a function $g(V)$ cannot be decreasing while maintaining a value above the diagonal $g(V) = V$. This leaves us with the first set of requirements, $g'(V) > 0$ and $g(V) < V$. The resulting space that g is allowed to span is charted in Fig. 3.

For matching, $V = \frac{R_1(E_1)}{E_1} = \frac{R_2(E_2)}{E_2} = \ldots = \frac{R_n(E_n)}{E_n}$ is the same for all options i. Thus, when $\frac{d^2R_i(E_i)}{dE_i^2} = g'(V)\left(\frac{g(V)-V}{E_i}\right)$ assumes a positive or a negative sign, it does so equally for all options, and the sign is governed solely by the properties of $g(V)$. For matching yielding a reward maximum ($g'(V) > 0$ and $g(V) < V$), it is then apparent that $\frac{d^2R_i(E_i)}{dE_i^2} < 0$ for all options and for any E_i. This proves that all options provide diminishing returns, and they do so for all values of effort.

Specific Examples. The previous section showed that matching provides a reward maximum when $g'(V) > 0$ and, for actual reward returns, $g(V) < V$. These requirements constrain the parameters of the solutions and therefore also dictate the nature of the solutions. For instance, in Eqs. **5** and **6** for

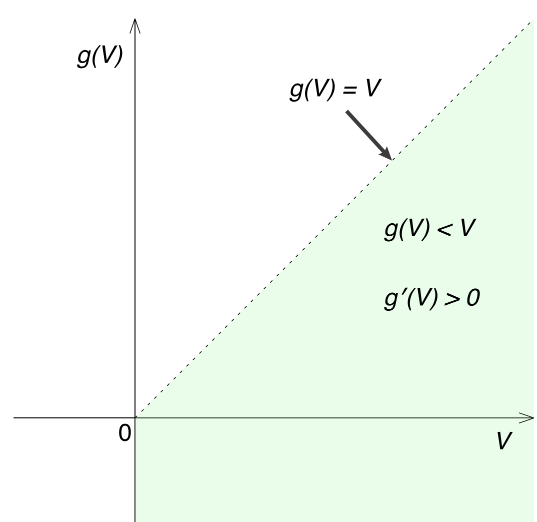

Fig. 3. Generation of solutions for which matching yields reward maxima. For matching to deliver a reward maximum, the generator function must have its derivative greater than zero ($g'(V) > 0$), and in addition, for actual reward return situations, it must be $g(V) < V$. Combined with the fact that for matching, $V = \frac{R}{E} > 0$, the space of the admissible $g(V)$ values is highlighted in green.

PSYCHOLOGICAL AND COGNITIVE SCIENCES

which $g(V) = aV$, $g'(V) > 0$ requires that $a > 0$. It then becomes obvious that a reward–effort function $a\hat{R}\ln(E)$ shows diminishing returns with increasing effort E. As another example, in Eq. **6**, in addition to $a > 0$ that again follows from $g'(V) > 0$, the requirement $g(V) < V$ dictates that $aV < V$ and so $a < 1$. It is apparent that a reward–effort function RE^a with $0 < a < 1$ also shows diminishing returns. And, in Eq. **7**, the $g(V)$ requirements on a reward maximum impose $a > 0$ and $c_i > 0$. It is easy to see that a reward–effort function $\frac{E}{a+c_i E}$ with $a, c_i > 0$ also shows diminishing returns.

The previous section proved that this is a general finding. For matching to deliver a reward maximum, the reward–effort contingencies of all choice options must show diminishing returns.

Increasing Returns. It is worth also considering the complement, i.e., the possibility that matching might be optimal under situations of increasing (and not diminishing) returns. Increasing returns for an option j imply $\frac{d^2 R_j(E_j)}{dE_j^2} = R_j(E_j)'' > 0$. *Reward Maxima and Diminishing Returns* showed that when Eq. **4** is satisfied and matching holds, $\frac{d^2 R_i(E_i)}{dE_i^2}$ exhibits the same sign for all options i and for all values of effort E_i. Under these conditions, it follows that when $\frac{d^2 R_j(E_j)}{dE_j^2} > 0$ for an option j, it must be $\frac{d^2 R_i(E_i)}{dE_i^2} = R_i(E_i)'' > 0$ for all options i. In this case, the principal minors of the bordered Hessian (*Optimal Decision Making*) become all negative. But negative principal minors constitute a sufficient condition for a reward minimum (34). Thus, increasing returns do not allow for matching to be optimal; diminishing returns are indeed required. An additional consequence of $\frac{d^2 R_i(E_i)}{dE_i^2}$ exhibiting the same sign for all options is that the leading principal minors (*Optimal Decision Making*) either alternate in sign—which is a sufficient condition for a reward maximum—or are all negative—which is a sufficient condition for a minimum. A third possible outcome—a saddle point—would require a distinct pattern of signs of the principal minors (34).

ACKNOWLEDGMENTS. I thank Drs. Leonard Green, Lawrence Snyder, Nina Miolane, and Julian Brown for helpful comments. This work was supported by the NIH Grant K99NS100986 and by the McDonnell Center for Systems Neuroscience.

1. Friedman M (1953) *Essays in Positive Economics* (Univ of Chicago Press, Chicago).
2. Becker G (1978) *The Economic Approach to Human Behavior* (Univ of Chicago Press, Chicago).
3. Mazur JE (1981) Optimization theory fails to predict performance of pigeons in a two-response situation. *Science* 214:823–825.
4. Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. *Econometrica* 47:263–291.
5. Körding K (2007) Decision theory: What "should" the nervous system do? *Science* 318:606–610.
6. Hollis M, Nell EJ (1975) *Rational Economic Man: A Philosophical Critique of Neo-Classical Economics* (Cambridge Univ Press, Cambridge, UK).
7. Herrnstein RJ (1990) Rational choice theory: Necessary but not sufficient. *Am Psychol* 45:356–367.
8. McDowell J (2013) On the theoretical and empirical status of the matching law and matching theory. *Psychol Bull* 139:1000–1028.
9. Herrnstein RJ (1961) Relative and absolute strength of response as a function of frequency of reinforcement. *J Exp Anal Behav* 4:267–272.
10. Staddon J (1964) Reinforcement as input: Cyclic variable-interval schedule. *Science* 145:410–412.
11. Herrnstein RJ (1970) On the law of effect. *J Exp Anal Behav* 13:243–266.
12. Baum WM (1974) Choice in free-ranging wild pigeons. *Science* 185:78–79.
13. Davison M, McCarthy D (1988) *The Matching Law: A Research Review*. (Lawrence Erlbaum Assoc., Mahwah, NJ).
14. Herrnstein RJ (2000) *The Matching Law: Papers in Psychology and Economics* (Harvard Univ Press, Cambridge, MA).
15. Sugrue LP, Corrado GS, Newsome WT (2004) Matching behavior and the representation of value in the parietal cortex. *Science* 304:1782–1787.
16. Loewenstein Y, Seung HS (2006) Operant matching is a generic outcome of synaptic plasticity based on the covariance between reward and neural activity. *Proc Natl Acad Sci USA* 103:15224–15229.
17. Rachlin H (1971) On the tautology of the matching law. *J Exp Anal Behav* 15:249–251.
18. Rachlin H, Battalio R, Kagel J, Green L (1981) Maximization theory in behavioral psychology. *Behav Brain Sci* 4:371–388.
19. Binmore K (1991) Rational choice theory: Sufficient but not necessary. *Am Psychol* 46:797–799.
20. Staddon JER (1980) *Limits to Action: The Allocation of Individual Behavior* (Academic, Cambridge, MA).
21. Rachlin H (1982) Economics of the matching law. *Quant Anal Behav* 2:347–374.
22. Rachlin H, Green L, Tormey B (1988) Is there a decisive test between matching and maximizing? *J Exp Anal Behav* 50:113–123.
23. Herrnstein RJ, Loveland DH (1975) Maximizing and matching on concurrent ratio schedules 1. *J Exp Anal Behav* 24:107–116.
24. Baum WM (1981) Optimization and the matching law as accounts of instrumental behavior. *J Exp Anal Behav* 36:387–403.
25. Otto AR, Taylor EG, Markman AB (2011) There are at least two kinds of probability matching: Evidence from a secondary task. *Cognition* 118:274–279.
26. Kubanek J, Snyder LH (2015) Matching behavior as a tradeoff between reward maximization and demands on neural computation. *F1000Res* 4:147.
27. Rachlin H (1978) A molar theory of reinforcement schedules. *J Exp Anal Behav* 30: 345–360.
28. Heyman GM (1982) Is time allocation unconditioned behavior? *Quant Anal Behav* 2:459–490.
29. Shephard RW, Färe R (1974) *The Law of Diminishing Returns in Production Theory* (Springer, Berlin), pp 287–318.
30. Brue SL (1993) Retrospectives: The law of diminishing returns. *J Econ Perspect* 7: 185–192.
31. Turgot ARJ (1999) The formation and distribution of wealth. Reflections on Capitalism, ed Jupp K (Othila Press, London).
32. Cobb CW, Douglas PH (1928) A theory of production. *Am Econ Rev* 18: 139–165.
33. Knight FH (1944) Diminishing returns from investment. *J Polit Econ* 52:26–47.
34. Baumol WJ (1977) *Economic Theory and Operations Analysis* (Prentice-Hall, Englewood Cliffs, NJ).
35. Jones C (2013) *Introduction to Economic Growth* (WW Norton & Co., New York).
36. Real LA (1980) On uncertainty and the law of diminishing returns in evolution and behavior. *Limits to Action: The Allocation Individual Behavior*, ed Staddon JER (Academic, Cambridge, MA), pp 37–64.
37. Fagen R (1987) A generalized habitat matching rule. *Evol Ecol* 1:5–10.
38. Morris DW (1994) Habitat matching: Alternatives and implications to populations and communities. *Evol Ecol* 8:387–406.
39. Charnov EL (1976) Optimal foraging, the marginal value theorem. *Theor Popul Biol* 9:129–136.
40. Myerson J, Green L (1995) Discounting of delayed rewards: Models of individual choice. *J Exp Anal Behav* 64:263–276.
41. Levy DJ, Glimcher PW (2012) The root of all value: A neural common currency for choice. *Curr Opin Neurobiol* 22:1027–1038.
42. Croxson PL, Walton ME, O'Reilly JX, Behrens TE, Rushworth MF (2009) Effort-based cost–benefit valuation and the human brain. *J Neurosci* 29:4531–4541.
43. Prévost C, Pessiglione M, Météreau E, Cléry-Melin ML, Dreher JC (2010) Separate valuation subsystems for delay and effort decision costs. *J Neurosci* 30:14080–14090.
44. Vaughan W (1981) Melioration, matching, and maximization. *J Exp Anal Behav* 36:141–149.
45. Körding KP, Wolpert DM (2006) Bayesian decision theory in sensorimotor control. *Trends Cognit Sci* 10:319–326.
46. Soltani A, Wang XJ (2010) Synaptic computation underlying probabilistic inference. *Nat Neurosci* 13:112–119.
47. Herrnstein RJ, Prelec D (1991) Melioration: A theory of distributed choice. *J Econ Perspect* 5:137–156.
48. Lau B, Glimcher PW (2005) Dynamic response-by-response models of matching behavior in rhesus monkeys. *J Exp Anal Behav* 84:555–579.
49. Corrado GS, Sugrue LP, Seung HS, Newsome WT (2005) Linear-nonlinear-Poisson models of primate choice dynamics. *J Exp Anal Behav* 84:581–617.
50. Soltani A, Wang XJ (2006) A biophysically based neural model of matching law behavior: Melioration by stochastic synapses. *J Neurosci* 26:3731–3744.
51. Herrnstein RJ, Prelec D (1992) *Melioration* (Russell Sage Foundation, New York).
52. Staddon J, Motheral S (1978) On matching and maximizing in operant choice experiments. *Psychol Rev* 85:436–444.

Supporting Information

Kubanek 10.1073/pnas.1703440114

SI Methods

Probabilistic Reward Outcomes. Rewards in natural settings are often stochastic. The stochasticity can manifest in two ways. Either the reward itself is delivered with a certain probability or the conversion of a subject's decision to a choice is probabilistic. In both cases, a decision i provides reward R_i with probability p_i. In what follows, it is shown that matching rests on diminishing returns even in situations in which rewards are delivered in a probabilistic way and subjects maximize the expected value of the reward.

Optimal decision making with probabilistic rewards. When rewards R_i are delivered with probabilities p_i, an optimal decision maker distributes her effort to maximize the expected value over the reward options,

$$E[R] = p_1 R_1(E_1) + p_2 R_2(E_2) + \ldots + p_n R_n(E_n), \quad \textbf{[S1]}$$

subject to the constraint that total effort of the decision maker is limited, $\sum_i E_i = E_{\max}$.

Using the method of Lagrange multipliers to solve this problem, the Lagrangian is in this case formulated as

$$\mathcal{L}(E_1, E_2, \ldots, E_n, \lambda) = p_1 R_1(E_1) + p_2 R_2(E_2) + \ldots + p_n R_n(E_n) + \lambda(E_{\max} - \sum_i E_i). \quad \textbf{[S2]}$$

Setting the partial derivatives to 0, we get

$$\frac{\partial \mathcal{L}}{\partial E_i} = p_i \frac{dR_i(E_i)}{dE_i} - \lambda = 0,$$

and so

$$p_i \frac{dR_i(E_i)}{dE_i} = \lambda$$

for a certain λ. This means that

$$p_1 \frac{dR_1(E_1)}{dE_1} = p_2 \frac{dR_2(E_2)}{dE_2} = \ldots = p_n \frac{dR_n(E_n)}{dE_n}. \quad \textbf{[S3]}$$

Eq. **S3** dictates how effort should be allocated to attain a critical point in the reward landscape. This can be a maximum, a minimum, or a saddle point. To obtain a reward maximum, the leading principal minors of the bordered Hessian matrix corresponding to Eq. **S2**, evaluated at critical points, must alternate in sign, with the first minor (of order 3) showing a positive sign (34). As in *Methods* in the main text, let us label $\frac{d^2 R_i(E_i)}{dE_i^2} = R_i''(E_i)$ at a certain critical point E_i. The bordered Hessian for Eq. **S2** is

$$H^B(E_1, E_2, \ldots, E_n) = \begin{bmatrix} 0 & -1 & -1 & \ldots & -1 \\ -1 & p_1 R_1''(E_1) & 0 & \ldots & 0 \\ -1 & 0 & p_2 R_2''(E_2) & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & 0 & p_n R_n''(E_n) \end{bmatrix}.$$

For two options, the leading principal minor of order 3 is equal to

$$\det H^B(E_1, E_2) = -(p_1 R_1''(E_1) + p_2 R_2''(E_2)).$$

For a critical point to represent a maximum, this leading principal minor must be positive, so it must be

$$p_1 R_1''(E_1) + p_2 R_2''(E_2) < 0.$$

Because $p_i > 0$, it is easy to see that the same two observations we made in the main text hold also in this case. First, for a critical point to represent a reward maximum—so that the above equation holds—at least for one option and at least for some value of E_i the second derivative $R_i''(E_i)$ must be $R_i''(E_i) < 0$. Second, if for any E_i and for both options $R_i''(E_i) < 0$, then the above equation always holds and a critical point is a maximum. In addition, as in the main text, these two observations hold for any number of options.

Matching Behavior. Matching can also be generalized to accommodate probabilistic reward, with subjects matching the expected values of rewards over the individual options:

$$\frac{p_1 R_1(E_1)}{E_1} = \frac{p_2 R_2(E_2)}{E_2} = \ldots = \frac{p_n R_n(E_n)}{E_n}. \quad \textbf{[S4]}$$

Matching–Optimization Relation. Analogously to choice situations in the main text, matching and optimization in the probabilistic reward settings align when

$$p \frac{dR}{dE} = g\left(\frac{pR}{E}\right). \quad \textbf{[S5]}$$

To see why, incorporating Eq. **S5** into Eq. **S3** leads to $p_1 \frac{dR_1(E_1)}{dE_1} = g\left(\frac{p_1 R_1(E_1)}{E_1}\right) = p_2 \frac{dR_2(E_2)}{dE_2} = g\left(\frac{p_2 R_2(E_2)}{E_2}\right) = \ldots = p_n \frac{dR_n(E_n)}{dE_n} = g\left(\frac{p_n R_n(E_n)}{E_n}\right)$. A strictly monotonic function is invertible. Inverting g, we obtain

$$\frac{p_1 R_1(E_1)}{E_1} = \frac{p_2 R_2(E_2)}{E_2} = \ldots = \frac{p_n R_n(E_n)}{E_n},$$

which represents the matching law in the expected value generalization (Eq. **S4**).

The contingencies between reward and effort, $R(E)$, for which these two models align, are merely scaled versions of the solutions provided in the main text, as shown next.

Reward–Effort Contingencies for Estimated Reward Returns. It is easy to see that for a simple case of $g(V) = aV$, Eq. **S5**, now becoming $p_i \frac{dR_i(E_i)}{dE_i} = a \frac{p_i \hat{R}_i}{E_i}$, yields the same reward–effort contingency as in the main text:

$$R_i(E_i) = a\hat{R}_i \ln(E_i) + c_i.$$

Reward–Effort Contingencies for Actual Reward Returns. Analogous findings are obtained for cases in which subjects match their response distribution to the actual returns obtained from choosing an option. For $g(V) = aV$, Eq. **S5**, now becoming $p_i \frac{dR_i(E_i)}{dE_i} = a \frac{p_i R_i(E_i)}{E_i}$, again yields the same solution as in the main text:

$$R_i(E_i) = R_i E_i^a.$$

The solutions become generalized forms of those provided in the main text when $g(V)$ is nonlinear. For example, for $g(V) = aV^2$, Eq. **S5** generates a reward–effort contingency

$$R_i(E_i) = \frac{E_i}{ap_i + c_i E_i}.$$

Note that in this case the probability p_i figures in the solution.

As in the main text, all other reward–effort contingencies can be generated by defining specific forms of $g(V)$. The admissible forms of this function, such that optimal decision making based on probabilistic rewards delivers maxima and not minima or saddle points, are provided next.

Reward Maxima and Diminishing Returns. It will be shown here that for matching to deliver a maximum in the expected reward when rewards are probabilistic, the reward–effort profiles of all choice options must show diminishing returns. As in the main text, the proof follows from evaluating the second derivative of $R(E)$.

Estimated reward returns. Applying the chain rule, the derivative of Eq. **S5** with respect to effort for estimated rewards $\hat{R}$ is equal to

$$\frac{d^2R(E)}{dE^2} = \frac{1}{p}g'\left(\frac{p\hat{R}}{E}\right)\left(-\frac{p\hat{R}}{E^2}\right).$$

As in the main text, *Optimal Decision Making* showed that for our constraint optimization problem to yield a reward maximum, it must be $\frac{d^2R(E)}{dE^2} < 0$ for at least one option and for at least some value of E. To meet that requirement, because in the above expression $p > 0$, $\frac{1}{p} > 0$, and $-\frac{p\hat{R}}{E^2} < 0$, it must be $g'(\frac{p\hat{R}}{E}) > 0$ for at least one option and for at least some E. But because of matching, the g' argument $\frac{p\hat{R}}{E}$ is equal for all options because $\frac{p_1\hat{R}_1}{E_1} = \frac{p_2\hat{R}_2}{E_2} = \ldots = \frac{p_n\hat{R}_n}{E_n} = V$. It follows that it must be $g'(V) > 0$, and this value is equal for all options. Thus, whether matching yields a reward maximum or a minimum depends solely on this requirement on g. Furthermore, when this requirement is fulfilled, it is apparent that $\frac{d^2R_i(E_i)}{dE_i^2} < 0$ for all options and for any values of E_i. This proves that the reward–effort contingencies for all options are negatively accelerated, i.e., show diminishing returns, and they do so for all values of effort.

Actual reward returns. The proof proceeds in a similar fashion for actually obtained reward returns $R(E)$. In this case, the derivative of Eq. **S5** with respect to effort is

$$\frac{d^2R(E)}{dE^2} = \frac{1}{p}g'\left(\frac{pR(E)}{E}\right)\left(p\frac{\frac{dR(E)}{dE}E - R(E)}{E^2}\right)$$
$$= \frac{1}{p}g'\left(\frac{pR(E)}{E}\right)\left(\frac{p\frac{dR(E)}{dE} - \frac{pR(E)}{E}}{E}\right).$$

Here, according to Eq. **S5**, $p\frac{dR(E)}{dE} = g\left(\frac{pR(E)}{E}\right)$. Thus,

$$\frac{d^2R(E)}{dE^2} = \frac{1}{p}g'\left(\frac{pR(E)}{E}\right)\left(\frac{g\left(\frac{pR(E)}{E}\right) - \frac{pR(E)}{E}}{E}\right).$$

The beauty of matching again shines through this equation. Given that $\frac{1}{p} > 0$ and $E > 0$, it is apparent that the sign of the right side of the equation is solely a function of the matching term $V = \frac{pR(E)}{E}$:

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

As in the main text, *Optimal Decision Making* showed that for a critical point to represent a reward maximum, it must be $\frac{d^2R(E)}{dE^2} < 0$ for at least one option and at least for some value of E. As shown in the main text, this requirement is met for $g'(V) > 0$ and $g(V) < V$.

For matching, $V = \frac{p_1R_1(E_1)}{E_1} = \frac{p_2R_2(E_2)}{E_2} = \ldots = \frac{p_nR_n(E_n)}{E_n}$ is the same for all options i. Thus, when $\frac{d^2R_i(E_i)}{dE_i^2} = \frac{1}{p_i}g'(V)\left(\frac{g(V)-V}{E_i}\right)$ assumes a positive or a negative sign, it does so equally for all options, and the sign is governed solely by the properties of $g(V)$. For matching yielding a reward maximum ($g'(V) > 0$ and $g(V) < V$), it is then apparent that $\frac{d^2R_i(E_i)}{dE_i^2} < 0$ for all options and for any E_i. This proves that all options provide diminishing returns, and they do so for all values of effort.

Together, all findings obtained for deterministic rewards in the main text also hold in situations in which reward is probabilistic and in which subjects operate on expected values of rewards.

Relationship to Prospect Theory. The formalism of maximizing the expected value of reward (Eq. **S1**),

$$E[R] = p_1R_1(E_1) + p_2R_2(E_2) + \ldots + p_nR_n(E_n),$$

is closely related to the criterion of prospect theory (4). In prospect theory, subjects maximize the expected utility in the form

$$E[U] = \pi(P_1)u(R_1(E_1)) + \pi(P_2)u(R_2(E_2)) + \ldots$$
$$+ \pi(P_n)u(R_n(E_n)),$$

where u is a utility function that operates on an outcome R_i, p_i is the probability that the associated reward is delivered, and π is a function that transforms probability into its subjective perception (4). When the function $R_i(E_i)$ represents the utility of effort E_i exerted at option i such that $R_i(E_i) = u(R_i(E_i))$, and when we realize that π is a function that is the same for each option and so $p_i = \pi(P_i)$, the two criteria are equal.

Note that in *SI Methods*, *Reward Maxima and Diminishing Returns*, the proof of matching resting on diminishing returns is very similar to that of the main text, with the exception of an additional multiplier $\frac{1}{p}$. This term is positive because $p > 0$, and so the sign consideration of the second derivative remains unchanged. In prospect theory, the $\frac{1}{p}$ multiplier merely becomes $\frac{1}{\pi(p)}$. This term is also positive because $\pi(p) > 0$ (4). This way, the prospect theory-based generalization provides the same conclusion. Thus, even in scenarios in which rewards are stochastic and in which people behave to maximize their expected utility, matching can be an optimal strategy as long as the choice environment features diminishing returns.