# Efficiency of learning vs. processing: Towards a normative theory of multitasking

**Yotam Sagiv (ysagiv@princeton.edu), Sebastian Musslick (musslick@princeton.edu),**
**Yael Niv (yael@princeton.edu), Jonathan D. Cohen (jdc@princeton.edu)**
Princeton Neuroscience Institute
Princeton University

## Abstract

A striking limitation of human cognition is our inability to execute some tasks simultaneously. Recent work suggests that such limitations can arise from a fundamental trade-off in network architectures that is driven by the sharing of representations between tasks: sharing promotes quicker learning, at the expense of interference while multitasking. From this perspective, multitasking failures might reflect a preference for learning efficiency over parallel processing capability. We explore this hypothesis by formulating an ideal Bayesian agent that maximizes expected reward by learning either shared or separate representations for a task set. We investigate the agent's behavior and show that over a large space of parameters the agent sacrifices long-run optimality (higher multitasking capacity) for short-term reward (faster learning). Furthermore, we construct a general mathematical framework in which rational choices between learning speed and processing efficiency can be examined for a variety of different task environments.

**Keywords:** multitasking; cognitive control; Bayesian inference; capacity constraints;

## Introduction

The human brain's ability to simultaneously perform distinct tasks contains a curious tension. On one hand, we are able to concurrently carry out a large number of actions (e.g. breathe, speak, chew gum, etc.) seemingly without exerting any effort. In contrast, some behaviors defy parallel execution (e.g. solving calculus problems and constructing shopping lists) and require serialization to successfully execute.

The distinction between sets of tasks that can be executed concurrently and those that cannot is often referred to in terms of a fundamental distinction between controlled and automatic processing (Posner & Snyder, 1975; Shiffrin & Schneider, 1977). Early theories attributed the inability to carry out multiple control-demanding tasks in parallel to reliance on a single, limited-capacity, serial processing mechanism – a hypothesis that has continued to dominate major theories of cognition (e.g., Anderson, 2013). The "multiple-resource hypothesis" presents a challenge to this view, arguing that multitasking limitations may reflect competition for the use of local resources (e.g., shared task-specific representations) by sets of tasks, rather than common reliance on a central control mechanism (Allport, Antonis, & Reynolds, 1972; Navon & Gopher, 1979; Meyer & Kieras, 1997). Under this view, the role of cognitive control is to resolve such conflicts when they arise by limiting processing to only a single task at a time (Cohen, Dunbar, & McClelland, 1990; Botvinick, Braver, Barch, Carter, & Cohen, 2001). That is, limiting processing is the *purpose of control*, rather than a reflection of a constraint on the control system itself. Recent computational work has provided a formal grounding for this argument, showing that even modest amounts of overlap between task representations

can drastically limit the number of tasks a network can engage at the same time without invoking interference among them (Feng, Schwemmer, Gershman, & Cohen, 2014; Musslick et al., 2016). Critically, this number appears to be relatively insensitive to the size of the network.

The findings above raise an important question: insofar as shared representation between tasks impose limitations on multitasking, why would a neural system prefer shared representations over separate ones? Insights into this question can be gained from the machine learning literature, where the learning of shared representations between tasks is considered a desirable outcome (Baxter, 1995; Caruana, 1998; Bengio, Courville, & Vincent, 2013). For instance, work on multi-task learning[1] suggests that shared representations between tasks promote faster learning, as well as better generalization performance across tasks (Caruana, 1997; Collobert & Weston, 2008). Moreover, learning dynamics in neural networks themselves promote the learning of shared representation based on shared structure in the task environment (Hinton, 1986; Saxe, McClelland, & Ganguli, 2013). Thus, there appears to be a fundamental trade-off in neural networks between the efficiency of learning (and generalization) on the one hand, and the efficiency of processing (i.e. multitasking capability) on the other hand (Musslick et al., 2017).

The trade-off between learning and processing efficiency constitutes an optimization problem that is dependent on the demands of the task environment. The work described here examines this optimization problem as a function of critical parameters, such as the differences in rate of learning for shared vs. separated representations, and the benefits gained by parallel over serial task performance. Analysis of this problem may help provide a formally rigorous, and even normative account of longstanding, well-characterized psychological phenomena, such as the common trajectory in skill acquisition from controlled to automatic processing (Shiffrin & Schneider, 1977; Logan, 1980).

Ideally, our analysis would build on formal characterization of the learning rate for different types of representations, given a specified learning algorithm (e.g. backpropagation). However, since this is not immediately available, to construct a probabilistic generative model we begin by assuming simple functional forms for the learning trajectory associated with shared vs. separated task representations in a multitasking environment, and then use the generative model to define

---

[1] Note that the term 'multi-task' differs from the term 'multitasking'. The former refers to the paradigm of training the same network on multiple tasks, whereas the latter refers to the process of carrying out multiple tasks concurrently.

an ideal Bayesian agent that behaves optimally inside that environment. Taken together, the environment and agent models provide a simple, normative framework in which questions about the learning-processing trade-off can be explored.

## A rational model of multitasking

We begin our analysis of the optimal balance between learning and processing efficiency by formalizing the task environment. We then describe how the agent model chooses between the use of shared vs. separate representations in that environment to optimize performance, which we define as maximizing reward over the entire horizon of performance.

### Task Environment

We consider an environment in which a task can be defined as a process (e.g. naming the color of a stimulus) that maps the dimension of a stimulus (e.g. color) to a particular response modality (e.g. verbal response). Here we assume that stimuli consist of $N$ dimensions (e.g. color, shape, and texture) and that responses are carried out over $K$ response modalities (e.g. naming, pointing, or looking), resulting in $NK$ possible tasks in any environment. We adopt a formal definition of multitasking from earlier work (Musslick et al., 2016; Alon et al., 2017), in which a multitasking condition is defined as the requirement to execute multiple tasks at the same time, none of which share a stimulus or response dimension. Consequently, at most $\min\{N, K\}$ tasks can be carried out concurrently.

The agent is asked to optimize performance over a series of $\tau$ multitasking trials. On each trial, the agent is asked to perform $\alpha$ tasks, where $\alpha$ is drawn from a latent multinomial distribution. We introduce multitasking pressure by specifying a reward schedule that favors concurrent performance of tasks. For every task answered correctly, the agent receives 1 unit of reward, resulting in $\alpha$ rewards if the agent is able to perform all tasks with maximal accuracy at the same time. However, if the agent chooses instead to perform all tasks sequentially, it loses $jC$ reward units on task $j$, where $j$ indexes the tasks from 0 to $\alpha - 1$ (so that the agent receives $\sum_{j=0}^{\alpha-1} 1 - jC$ rewards given maximal accuracy). $C$ is termed the "serialization cost" or "time cost". The per-task loss is linear in time taken, making the per-trial (cumulative) loss over all assigned tasks quadratic.

Optimization is defined as the choice, on each trial, of a performance strategy that maximizes total future reward; that is, summed over the current trial and the potentially discounted reward anticipated for each future trial. This requires estimating and convolving the expected multitasking requirements over trials, performance for executing the tasks concurrently vs. individually as a function of the estimated learning rate for each (see below), and the serialization costs associated with performing tasks sequentially.

### Agent

The agent is considered to be a rational decision-maker that chooses between two independent, trainable processing strategies that result from two extremes of how multiple tasks

can be represented in a single network. The first representation strategy is as a minimal basis set, in which all tasks relying on the same stimulus dimension encode the stimuli using the same (shared) set of hidden representations (i.e. $N$ sets of hidden representations) that are then mapped to the output dimensions for each of the tasks. The second strategy uses tensor product representations, in which each task encodes its stimuli using its own set of (separated) hidden representations (resulting in $NK$ sets of hidden representations) that are mapped to the output dimension for the task. While the minimal basis set provides a more efficient encoding of the stimuli, it does not permit multitasking since the use of shared representations introduces crosstalk between any pair of simultaneously activated tasks (see Figure 1; Feng et al., 2014; Musslick et al., 2016; Alon et al., 2017). Thus, use of the minimal basis set forces a serialization cost of $jC$ reward units for task $j = 1, 2, \ldots, \alpha - 1$. Conversely, the tensor product representation permits multitasking without interference, since each task is assigned its own set of hidden representations that comprise independent processing pathways in the network. We assume that the agent has the potential to develop both forms of representation, but these must be learned.
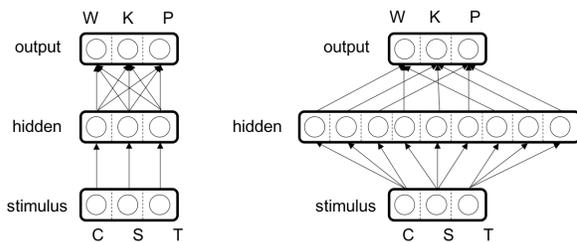


(a) maximal sharing ("basis")    (b) maximal multitasking capability ("tensor")

Figure 1: Schematic of network schemes that maximize representation overlap (a) vs. multitasking capability (b). C, S, T designate the stimulus dimensions ("color", "shape", and "texture"), while W, K, P designate the output modalities ("word", "keyboard", "point"). The hidden-layer representation of the stimulus in (a) is shared for all three tasks involving the same input dimension (minimal basis set representation), whereas in (b) a separate hidden-layer representation is dedicated to each task (tensor product representation).

Previous work has shown that, for a set of tasks that are in principle multitaskable, training using shared representations (such as a minimal basis set) leads to faster acquisition than learning separate representations for each task (such as a tensor product), as the former enables the sharing of learning signals across tasks (Musslick et al., 2017). We implement these effects by assuming that 1) the agent learns these two types of representations (i.e. processing strategies) by selecting and executing one or the other on each trial; 2) performance for each strategy improves as a function of the number of trials selected, and 3) learning is faster for the minimal basis strategy than for the tensor product strategy, as described below.

To model the learning of tasks, we define a probability of success function (aka "training function") for each of the two processing strategies. Let $f_B, f_T : \mathbb{N}_{\geq 0} \to [0, 1]$ denote these training functions for the minimal basis set and tensor product strategies, respectively. These serve as explicit characterizations of the agent's learning dynamics; $f_X(t)$ implements the learning curve by evaluating the probability of success on a given task after representation $X$ has been selected $t$ times. That is, every time the agent chooses to process the tasks in the trial using strategy $X$, the success probability for the task under strategy $X$ increases for the next time step. More formally, let $x_1, x_2, \ldots, x_n$ be a sequence of $n$ choices of representation. We define the probability that an agent succeeds when employing strategy $X$ on a task in trial $t$ as:

$$\mathbb{P}_X(\text{success on a task in trial } t) = f_X\Big(\sum_{i=0}^{t-1} \mathbb{1}_{x_i = X}\Big)$$

For convenience, we use the logistic function $f_X(t \mid k, t_0) = \frac{1}{1 + e^{-k(t - t_0)}}$. However, our analysis applies to any learning function that is monotonically increasing and bounded between $f_X(0) \approx 0$ and $\lim_{t \to \infty} f_X(t) = 1$. As noted above, we assume that learning occurs at a faster rate for the minimal basis set strategy as compared to tensor product strategy, and examine the influence of this difference by exploring a range of values for $k$, $t_0$ that together determine the rate of learning.

The agent uses standard Bayesian machinery to infer the expected reward given each representation, and then selects the representation that maximizes total discounted future reward. Specifically, let $\mathbb{E}_X[R]$ denote the expected reward for strategy $X$, $\mathbb{E}_X[R|t]$ denote the expected reward on trial $t$, and $\mu(t)$ be the temporal discounting function. Then we have that $\mathbb{E}_X[R] = \sum_{t=0}^{\tau} \mu(t) \mathbb{E}_X[R|t]$. Though temporal discounting can be irrational in many contexts, we note that a fully rational agent can be achieved with $\mu(t) = 1$.

Recall that $\alpha$ is the randomly assigned number of tasks required to be performed on a given trial. By marginalizing over $\alpha$, we get that the expected reward on each individual trial is $\mathbb{E}_X[R|t] = \sum_{i=1}^{\min\{N,K\}} \mathbb{P}(\alpha = i) \mathbb{E}_X[R|t, \alpha = i]$. Thus, the expected rewards for the minimal basis set and tensor product strategies correspond to

$$
\begin{aligned}
\mathbb{E}_B[R|t] &= \sum_{i=1}^{\min\{N,K\}} \mathbb{P}(\alpha = i) \sum_{j=0}^{i-1} \mathbb{P}_B(\text{success})(1 - jC) \\
\mathbb{E}_T[R|t] &= \sum_{i=1}^{\min\{N,K\}} \mathbb{P}(\alpha = i) \sum_{j=0}^{i-1} \mathbb{P}_T(\text{success})(1)
\end{aligned}
\tag{1}
$$

In order to compute the expected reward terms in Equation (1), the agent must be able to evaluate $\mathbb{P}(\alpha = i)$ and $\mathbb{P}_X(\text{success})$ by inferring the multinomial task distribution, as well as the training function $f_X$. The first can be inferred using Bayes' theorem, by keeping track of the number of times each particular $\alpha$ value was seen, in conjunction with a Dirichlet prior (we start from a uniform prior, implying absence of strong *a priori* belief about the distribution).

Inferring the parameters for the two training functions $f_B, f_T$ can similarly be done by tracking the history of successes and failures and then performing a Bayesian logistic

regression (intuitively, this can be understood as the agent inferring how fast it will learn). In this model, $k$ and $t_0$ have independent normal priors centered on their true values with high variance. Finally, we assume that the agent already knows $\tau$, the sequential processing cost $C$, and the temporal discounting function $\mu(t)$.

Once the expected values are computed, the agent must select an action. We assume this is done using a standard explore-exploit algorithm, the $\varepsilon$-greedy rule, in which the agent picks the action associated with greatest value with probability $1 - \varepsilon$, and uniformly otherwise.

## Formal analysis of equilibrium

We begin by analyzing an agent that has perfect knowledge about the task environment and learning rate, in order to assess performance independently of noise that might be generated by an inference process over these factors. This allows us to analytically derive equilibrium conditions under which the agent should be indifferent between the minimal basis set and the tensor product strategies. For this section, we let $N < K$ so that $N = \min\{N, K\}$ without loss of generality.

Observe that the expressions in Equation (1) reduce to:

$$
\begin{aligned}
\mathbb{E}_B[R|t] &= f_B(t) \mathbb{E}[g(\alpha, C)] \\
\mathbb{E}_T[R|t] &= f_T(t) \mathbb{E}[\alpha]
\end{aligned}
\tag{2}
$$

where $g(i, C) = \sum_{j=0}^{i-1}(1 - jC)$. Note that $g(i, C)$ encodes the amount of reward accrued by the agent for completing $i$ tasks in a serial fashion with time cost $C$. Plugging Equation (2) into the expression for the expected reward of both strategies we can express the condition for which the agent should be indifferent between them:

$$\frac{\mathbb{E}[\alpha]}{\mathbb{E}[g(\alpha, C)]} = \frac{\sum_{t=0}^{\tau} \mu(t) f_B(t)}{\sum_{t=0}^{\tau} \mu(t) f_T(t)} \tag{3}$$

An interesting property of this result is that agent-related and environmental parameters are analytically separable. Observe that the expectation terms on the left correspond to the agent's expected reward at asymptotic performance levels, and that the sum terms on the right denote the number of expected successes in a critical time period specified by the conjunction of the temporal discounting function and the training function. The indifference point can be understood intuitively as a surface over which the ratio of expected eventual rewards is equal to the ratio of times at which they are likely to be accrued (discounted by time). That is, the left side contains the ratio of the rewards the agent expects to earn if it is always correct, whereas the right side is a ratio of functions that weight when the agent prefers to receive the rewards.

Recall that $\mathbb{E}[g(\alpha, C)]$ corresponds to $\mathbb{E}[\sum_{j=0}^{i-1}(1 - jC)] = \mathbb{E}\left[\frac{\alpha}{2}\left(1 + [1 - (\alpha - 1)C]\right)\right]$. Since $C$ is a constant, it can be isolated from the expectation in Equation (3) to get an expression for the precise value of the serialization cost that characterizes the indifference surface. That is:

$$C_{eq} = \frac{2\mathbb{E}[\alpha]\left(1 - \frac{\sum_{t=0}^{\tau} \mu(t) f_T(t)}{\sum_{t=0}^{\tau} \mu(t) f_B(t)}\right)}{\mathbb{E}[\alpha(\alpha - 1)]} \tag{4}$$

Equation (4) provides a rigorous characterization of the trade-off between basis set and tensor product learning in multitasking environments described in the Introduction:

1. As the average number of parallel tasks increases, the cost of serialization must vanish for minimal basis set representations to remain preferable: $\mathbb{E}[\alpha] \to \infty \implies C_{eq} \to 0$.

2. As the learning benefit of shared representations diminishes, the value of shared representations disappears. That is, as the ratio between the (discounted) tensor product and basis set training functions approaches unity, for the latter to remain preferable the cost of serialization must tend toward zero: $\frac{\sum_{t=0}^{\tau}\mu(t)f_T(t)}{\sum_{t=0}^{\tau}\mu(t)f_B(t)} \to 1 \implies C_{eq} \to 0$.

3. $\frac{\sum_{t=0}^{\tau}\mu(t)f_T(t)}{\sum_{t=0}^{\tau}\mu(t)f_B(t)} \to 0 \implies C_{eq} \to \frac{2\mathbb{E}[\alpha]}{\mathbb{E}[\alpha(\alpha-1)]}$: As the ratio of the discounted training functions for the tensor product and minimal basis set representations approaches 0, the equilibrium-defining serialization cost becomes a function of the number of tasks required to be performed. Particularly, $C_{eq}$ is the serialization cost that sets expected reward for the minimal basis set representation to 0. This implication is not immediately obvious. Consider the task distribution $\mathbb{P}[\alpha = 1] = \mathbb{P}[\alpha = 2] = 1/2$. In this environment, $C_{eq} = 3$ and at asymptotic performance levels, the agent expects to win 1 reward unit when $\alpha = 1$, or win $-1$ when $\alpha = 2$. This makes sense; if learning tensor product representations is so much slower than minimal basis set representations that the ratio of the sums goes to 0, the agent is indifferent only if the expected earnings are 0.

Finally, we note that we have used arbitrary reward functions for the analyses above. However, it is possible to generalize the equilibrium condition in Equation (3) to any stationary reward function (i.e. does not change over the course of the experiment). Let $g_B(\alpha, j, C)$ denote a reward function with arbitrary dependence on the number of tasks currently being executed $\alpha$, the index of the task currently being executed $j$, or the serialization cost $C$; specifically, $g_B$ is the reward function used when the tasks are being executed serially. Furthermore, let $h_B(i, C)$ be the total reward gathered when $g_B$ is applied to each of the $i$ assigned tasks so that $h_B(i, C) = \sum_{j=0}^{i-1} g_B(i, j, C)$. Finally, define $g_T, h_T$ analogously for the case the tasks are being processed concurrently. Then a generalized equilibrium condition is:

$$\frac{\mathbb{E}[h_T(\alpha, C)]}{\mathbb{E}[h_B(\alpha, C)]} = \frac{\sum_{t=0}^{\tau}\mu(t)f_B(t)}{\sum_{t=0}^{\tau}\mu(t)f_T(t)} \qquad (5)$$

Observe that for $g_B = 1 - jC$ and $g_T = 1$, $h_B = g$ and $h_T = \alpha$ from Equation (3). The existence of this generalized equilibrium condition allows a large set of questions to be phrased within this framework. For example, it is easy to include an explicit cost of cognitive control (e.g. Shenhav et al., 2017) by adding a term to the basis set reward function that implements a cost that increases with the number of tasks executed.

## Numerical analysis with parameter inference

The analysis above characterized the behavior of an agent with perfect knowledge of the task environment and its learning functions. Here we relax these assumptions, and use numerical simulations[2] to evaluate the behavior of an agent that must infer these parameters. We assess the agent's performance across a series of task environments and learning specifications by crossing a set of reasonable parameter ranges.

We let $\tau = 1000$. We set $C \in [0, 1]$, varying from no punishment to receiving no reward for a correct answer. We use an exponential discounting scheme $\mu(t) = \gamma^{-0.025t}$ for $\gamma \in [0.5, 1.0]$. This covers the range from extreme discounting to no discounting at all. We characterize the training functions as logistic with $f_X(t) = \frac{1}{1+e^{-0.1(t-t_X)}}$. This allows us to precisely characterize difference in learning rates through the ratio $t_T/t_B$. To that end, we set $t_B = 200$, reflecting the speed of minimal basis set learning, and let $t_T$ vary in $[200, 600]$. We let $N = K = 4$ and define the distribution over tasks as $\mathbb{P}(\alpha = 1) = 0.7$, $\mathbb{P}(\alpha = 2) = \mathbb{P}(\alpha = 3) = \mathbb{P}(\alpha = 4) = 0.1$ so that the intensity and frequency of multitasking trials is sufficient to permit either strategy given appropriate parameters. We set $\varepsilon = 0.1$ to facilitate early exploration of the tensor product option in the face of immediate rewards due to the minimal basis set option. Finally, we quantify the agent's strategy preference as $\mathbb{P}(\text{pick } X) = \frac{\text{number of times } X \text{ was picked}}{\tau}$, and track how $\mathbb{P}(\text{pick basis set})$ varies with the parameters[3].
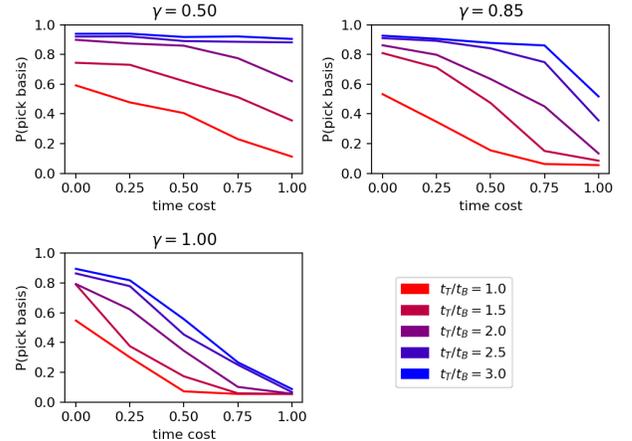
Figure 2: Simulation results for the inference model. $t_T/t_B$ refers to the midpoint ratio of the tensor product and minimal basis set training functions. Time cost denotes the value of $C$. Note that the agent increases their preference for the minimal basis set representation when the time cost is decreased, the learning rate ratio is increased, or gamma is decreased.

The results (see Fig. 2) show that there is a broad range of parameterizations under which the agent will opt for selecting the minimal basis set strategy over the tensor product strategy ($\mathbb{P}(\text{pick basis set}) > 0.5$). These preferences align with the normative analysis of how the parameters should affect overall preference: preference for the mini-

---

[3]We can use Equation (4) to show that even with weak discounting ($\gamma = 0.90$) and a modest learning rate ratio $t_T/t_B = 2$, the importance of fast training is such that the time cost must nearly equal the reward value ($C_{eq} \approx 0.75$) for indifference in this environment.

mal basis set strategy increases with relative speed of learning, decreases with serialization cost, and increases with the strength of temporal discounting as indicated by the linear model fit $\mathbb{P}(\text{select basis set}) \sim b_1 \times \frac{t_T}{t_B} + b_2 \times \text{timeCost} + b_3 \times \gamma$ ($b_1 = 0.25, t(78) = 47.26, p < 0.001; b_2 = -0.52, t(78) = -49.38, p < 0.001; b_3 = -0.64, t(78) = -35.78, p < 0.001$).

## Discussion

The constraints on human multitasking abilities present an interesting puzzle given the enormous processing capability of the brain. Here, we explored the hypothesis that this reflects a fundamental trade-off between learning and processing efficiency (Musslick et al., 2017), in which preference for learning to perform a set of tasks faster, which relies on the use of shared representations (Caruana, 1998; Baxter, 1995), comes at the expense of multitasking efficiency (Allport et al., 1972; Navon & Gopher, 1979; Meyer & Kieras, 1997; Feng et al., 2014; Musslick et al., 2016). This trade-off between the value of shared vs. separated representations is reminiscent of the complementary learning systems hypothesis (McClelland, McNaughton, & O'Reilly, 1995), which proposes the existence of two independent learning mechanisms. The first relies on shared representations to support inference, and the second uses separate representations to avoid the cost of catastrophic interference for memory encoding and retrieval. Thus, the trade-off between shared and separated representations appears to a fundamental one, that has different consequences in different processing contexts. Here, we have provided a normative analysis of this trade-off in the context of task performance that, under various assumptions, defines the conditions under which limitations in multitasking ability can be viewed as a result of optimal decision-making.

Agent behavior in our model was governed by several factors: the distribution of multitasking opportunities within the environment, the cost of serial vs. parallel performance, the rate at which each strategy is learned, and the discount rate for future rewards. The broad range of these factors over which the minimal basis set strategy was optimal suggests that the theory provides a plausible account of why so many skills (e.g. driving a car, playing an instrument) seem to rely on cognitive control and serial execution during acquisition.

Theories of bounded rationality (Simon, 1955, 1982; Gigerenzer, 2008) assume that suboptimalities in human behavior arise from the use of heuristics rather than full deliberation, given the bounds of limited multitasking capacity and limited available information. Research in artificial intelligence has suggested that such behavior is normative; that is, it may reflect bounded *optimality*, in which an agent maximizes reward per unit time given intrinsic limitations in its computational architecture (Russell & Subramanian, 1995). The principles of bounded optimality are reflected in psychological models of cognition, in which humans perform optimally within the constraints of the cognitive system (Griffiths, Lieder, & Goodman, 2015; Gershman, Horvitz, & Tenenbaum, 2015). Yet, these accounts do not explain why computational limitations exist in the first place, other than the assumption of limited processing power/speed. The work here suggest that the bounds may arise from a normative response to constraints imposed by trade-offs intrinsic to *any network architecture*, whether neural or artificial – specifically, the trade-off between the advantages of faster learning and generalization provided by shared representations, and the advantages of concurrent parallelism and processing efficiency provided by separated representations (Musslick et al., 2017). Under this framework the source of the limitation is not in the brain/computing device, but rather in the fact that time in life is finite (i.e. the benefits of learning a task quickly far outweigh the value of learning it "optimally").

Of course, the model we described is relatively simple, and can be extended in a number of ways. Rather than using a logistic function to characterize learning, it may be more reasonable to scale the benefit of shared representations by the number of tasks (e.g. as in Musslick et al., 2017), or to implement the learning dynamics of actual neural networks on similar task spaces. Additionally, a cost of control parameter could be incorporated that scales with the number of tasks being executed and/or the complexity of the task environment (Shenhav, Botvinick, & Cohen, 2013). It is also plausible to consider the transfer of learning between the two strategies (i.e. generalization). This may be an important factor in shaping how representations evolve from the minimal basis set to tensor product forms over the course of training, as suggested by some neural evidence (Garner & Dux, 2015).

One might also consider meta-learning. The simulated agents learned about their task environment and learning functions, but always began with the same predetermined, static priors. It is possible that repeated experience over different task domains could inform these priors, improving the initial estimates of the learning functions. This would induce a higher rate of convergence to the optimal decision for cases in which the agent's prior experiences are relevant, and might also explain any reluctance to switch away from suboptimal decision-making in contexts where its experience is misleading. Such effects could be informative to similar lines of inquiry regarding separate mechanisms for goal-directed and habitual responding in mammals undergoing instrumental conditioning (Yin & Knowlton, 2006).

In sum, the results presented here strongly support the proposal that constraints in multitasking observed in human performance may arise from a normative approach to an inescapable trade-off between the value of rapidly acquiring a set of novel skills, and optimizing the efficiency with which these skills can be exercised. Such a normative theory of multitasking may have value not only for understanding human performance, but also for the design of artificial systems. Having a formal language with which to consider the trade-off between learning efficiency and multitasking capability (and the closely related constructs of controlled vs. automatic processing) will facilitate precise analysis of the design of autonomous agents that are capable not only of guiding their own actions, but also of learning the best ways of doing so.

# References

Allport, A., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology*, *24*(2), 225-235. doi: 10.1080/00335557243000102

Alon, N., Reichman, D., Shinkar, I., Wagner, T., Musslick, S., Cohen, J. D., . . . Ozcimder, K. (2017). A graph-theoretic approach to multitasking. In *Advances in neural information processing systems* (pp. 2097–2106).

Anderson, J. R. (2013). *The architecture of cognition*. Psychology Press.

Baxter, J. (1995). Learning internal representations. In *Proceedings of the eighth annual conference on computational learning theory* (pp. 311–320).

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624.

Caruana, R. (1997). Multitask learning. *Machine learning*, *28*(1), 41–75.

Caruana, R. (1998). Multitask learning. In S. Thrun & L. Pratt (Eds.), *Learning to learn* (pp. 95–133). Boston, MA: Springer US. doi: 10.1007/978-1-4615-5529-2_5

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing model of the stroop effect. *Psychological Review*, *97*.

Collobert, R., & Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on machine learning* (pp. 160–167).

Feng, S., Schwemmer, M., Gershman, S. J., & Cohen, J. D. (2014, 01). Multitasking versus multiplexing: Toward a normative account of limitations in the simultaneous execution of control-demanding behaviors. *Cognitive, affective & behavioral neuroscience*, *14*.

Garner, K., & Dux, P. E. (2015, 10). Training conquers multitasking costs by dividing task representations in the frontoparietal- subcortical system. *Proceedings of the National Academy of Sciences*.

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278.

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on psychological science*, *3*(1), 20–29.

Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in cognitive science*, *7*(2), 217–229.

Hinton, G. E. (1986). Learning distributed representations of concepts. In *Proceedings of the 8th confererence of the Cognitive Science Society* (pp. 1–12). Hillsdale, NJ: Lawrence Erlbaum Associates.

Logan, G. D. (1980). Attention and automaticity in stroop and priming tasks: Theory and data. *Cognitive psychology*, *12*, 523-53.

McClelland, J., McNaughton, B., & O'Reilly, R. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*.

Meyer, D., & Kieras, D. (1997, 02). A computational theory of executive cognitive processes and multiple-task performance: Part 1. basic mechanisms. *Psychological Review*, *104*, 3-65.

Musslick, S., Dey, B., Özcimder, K., Mostofa, M., Patwary, A., Willke, T., & Cohen, J. D. (2016, 08). Controlled vs. automatic processing: A graph-theoretic approach to the analysis of serial vs. parallel processing in neural network architectures. In *Proceedings of the 38th annual conference of the Cognitive Science Society* (pp. 1547–1552).

Musslick, S., Saxe, A., Özcimder, K., Dey, B., Henselman, G., & Cohen, J. D. (2017, August). Multitasking capability versus learning efficiency in neural network architectures. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (p. 829-834).

Navon, D., & Gopher, D. (1979). On the economy of the human-processing system. *Psychological Review*, *86*.

Posner, M., & Snyder, C. (1975). attention and cognitive control. In *Information processing and cognition: The loyola symposium* (pp. 55–85).

Russell, S. J., & Subramanian, D. (1995). Provably bounded-optimal agents. *Journal of Artificial Intelligence Research*, *2*, 575–609.

Saxe, A. M., McClelland, J. L., & Ganguli, S. (2013). Learning hierarchical category structure in deep neural networks. In *Proceedings of the 35th annual meeting of the cognitive science society* (pp. 1271–1276).

Shenhav, A., Botvinick, M., & Cohen, J. D. (2013, 07). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, *79*, 217-40.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., L Griffiths, T., D Cohen, J., & Botvinick, M. (2017, 01). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, *40*.

Shiffrin, R., & Schneider, W. (1977, 03). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*, 127-190.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, *69*(1), 99–118.

Simon, H. A. (1982). *Models of bounded rationality. 1982.* Cambridge: MIT Press.

Yin, H., & Knowlton, B. (2006). The role of the basal ganglia in habit formation. *Nature reviews. Neuroscience*, *7*.