INSEAD

The Business School for the World®

Working Paper Series

2019/02/TOM

High-Performance Practice Processes

Guillaume Roels

INSEAD, guillaume.roels@insead.edu

Forthcoming at *Management Science*

Despite their idiosyncrasies, motor and cognitive learning and endurance sports training have in common that they involve repeated practice. While considerable research has been devoted to the effect of practice on performance, little is known about optimal practice strategies. In this paper, we model the practice process for both skill acquisition and retention and optimize its profile to maximize performance on a predefined date. For skill acquisition, we find that the optimal process involves multiple phases of practice increase and decrease, yielding U-shaped effort consistent with the principle of distributing practice, and that the transitions between phases are smoother for skills that are easily forgotten (e.g., cognitive skills) than for those that are easily retained (e.g., continuous motor skills). In particular for the latter, an extended period of rest should precede an ultimate high-intensity stress. For skill retention, the optimal practice strategy consists of cycles of either constant effort (for skills that are easily forgotten) or pulsed effort (for skills that are easily retained) consistent with the principle of alternating stress and rest. Our parametric model thus indicates when commonly used high-performance practice strategies are indeed optimal.

Keywords: Process Optimization; People-centric Operations; Learning; Training; Education; Sports Analytics

High-Performance Practice Processes

Guillaume Roels
INSEAD, Boulevard de Constance, 77305 Fontainebleau, France, guillaume.roels@insead.edu

Despite their idiosyncrasies, motor and cognitive learning and endurance sports training have in common that they involve repeated practice. While considerable research has been devoted to the effect of practice on performance, little is known about optimal practice strategies. In this paper, we model the practice process for both skill acquisition and retention and optimize its profile to maximize performance on a predefined date. For skill acquisition, we find that the optimal process involves multiple *phases* of practice increase and decrease, yielding U-shaped effort consistent with the principle of distributing practice, and that the transitions between phases are smoother for skills that are easily forgotten (e.g., cognitive skills) than for those that are easily retained (e.g., continuous motor skills). In particular for the latter, an extended period of rest should precede an ultimate high-intensity stress. For skill retention, the optimal practice strategy consists of *cycles* of either constant effort (for skills that are easily forgotten) or pulsed effort (for skills that are easily retained) consistent with the principle of alternating stress and rest. Our parametric model thus indicates when commonly used high-performance practice strategies are indeed optimal.

Key words: process optimization, people-centric operations, learning, training, education, sports analytics
History:

1. Introduction

Consider the following three individuals: a runner training for a marathon, a machine operator learning to operate a new machine, and a prospective MBA student preparing for a GMAT test. Although they pursue different goals, they each go through a *practice process*, with a specific performance horizon.[1] Consistent with the saying "practice makes better" and the "10,000-hour rule" popularized by Gladwell (2008), the link between deliberate practice and performance is well established (Ericsson et al. 1993). However, practice *per se* may not suffice to improve performance. Numerous scientific studies on the effect of practice on performance indeed indicate that, among other factors, the type and timing of practice matter.

Building on these scientific studies, we take an optimization perspective on the practice process and address the questions 'What is the optimal practice process to maximize performance on a given date? Should the practice involve breaks? Should it be massed or distributed? Should it be pulsed? Should there be a decrease in practice intensity before the final performance test? How is it affected by the stage of skill development (acquisition vs. retention), the type of skill (cognitive vs. motor), and the difficulty of the final performance assessment?' We consider a setting

[1] Although physical training, motor learning, and cognitive learning involve different functions, the effect of practice on performance often relies on similar behavioral mechanisms (Schmidt and Bjork 1992). As Stulberg and Magness (2017, p. 21) put it, "Whether someone is trying to qualify for the Olympics, break ground in mathematical theory, or craft an artistic masterpiece, many principles underlying healthy, sustainable success are the same."

where practices are free and unconstrained (unlike, say, a course with pre-defined session duration) and individuals are captive learners (e.g., no drop-outs). We focus on optimizing the distribution (intensity and timing) of practice and ignore other on-task and off-task practice considerations (e.g., variability in the content of practice, motivation; see Anderson 2000).

Practice process optimization offers business opportunities for sports analytics, training in organizations, and business model innovations in education, as set out below.

The sport industry, which is worth \$480-\$620 billion globally (Collignon et al. 2011), has widely adopted analytics to optimize athletes' performance, as popularized by the books *Moneyball* (Lewis 2004) and *Soccernomics* (Kuper and Szymanski 2014), with direct implications for advertising (e.g., Nike's Breaking2 two-hour marathon) and entertainment (e.g., broadcasting revenue). Moreover, the highly visible successes of analytics in sports have convinced business leaders to adopt a more analytical approach to management, e.g., for performance improvement (Davenport 2014).

In business, many operations (e.g., agriculture, factory work, call center operations) involve routine tasks for which performance can be improved through repeated practice. Similar to an organization's learning curve (Wright 1936), there are numerous opportunities for optimizing learning at the individual level, in the spirit of Taylor's time-and-motion studies (Taylor 1911). Companies regularly require their workforce to undergo training (e.g., pilots use flight simulators), certification programs (e.g., 6 sigma), rehearsals (e.g., plays, concerts), and drills (e.g., business continuity) to improve their performance (Fisher et al. 2018). Employees themselves often engage in deliberate practice to improve their performance (Sonnentag and Kleine 2000). In fact, lifetime learning has become an economic imperative as routine tasks are increasingly automated or offshored due to advances in technology, forcing people to reorient their careers (*The Economist* 2017, Staats 2018).

Practice process optimization could also lead to business model innovations in education. Global education, which is worth \$4.4 trillion (Strauss 2013), faces rising costs and disruption by information and communication technologies that allow customization of course content and pacing (Terwiesch and Ulrich 2014). For instance, SwissVBS's ECHO's learning app incorporates findings from scientific research on cognitive learning to customize learning and optimize retention.

Our behavioral model builds upon the discrete-time, finite-horizon *fitness-fatigue* model proposed by Banister et al. (1975). Although this model was developed in the context of endurance sports training, practice process optimization can be applied to other contexts provided that the effect of practice on performance is well understood. In this model, each practice session has a positive impact on performance by contributing to a "stock of fitness" and a negative impact by contributing to a "stock of fatigue." Between practice sessions, both stocks decay exponentially. Thus performance results from balancing stress and rest, consistent with a popular framework for peak performance in sports, science, and the arts (Stulberg and Magness 2017).

Although the fitness-fatigue model aptly describes performance in endurance sports (Taha and Thomas 2003), it is not without limitations when it comes to optimization. Specifically, as the model is linear, optimizing the profile of the practice process results in a corner solution, hence the optimal solution is heavily dependent on specifying its feasible set. Moreover, it is typically a threshold policy, e.g., it prescribes a phase of high-intensity practice followed by a phase of low-intensity practice, in stark contrast with empirical evidence of the effectiveness of distributed practice. Furthermore, because the performance objective function is separable, practices are independent of each other and thus unaffected by the difficulty of the final performance assessment.

To overcome these shortcomings, we modify the fitness-fatigue model in three ways. First, we incorporate the idea that *adaptation* plays a central role in practice processes. We model the impact of a practice session on performance as a function of the *effort* expended, measured as the current practice intensity relative to a base level, corresponding to the mean intensity of past practice sessions. We consider two models of adaptation: (i) the geometric mean, which adapts slowly to high-intensity practice but quickly to low-intensity practice, as is typically the case for skills that are easily forgotten, such as *cognitive skills*; and (ii) the maximum mean, which adapts quickly to high-intensity practice but slowly to low-intensity practice, as is typically the case for skills that are retained over a long time horizon, such as *continuous motor skills* (e.g., skiing, riding a bicycle). Second, we introduce *nonlinearities* to capture the diminishing marginal returns of practice on fitness and its increasing marginal costs in terms of fatigue. Third, we propose a *multiplicative* model of performance for *skill retention* as a complement to the original *additive* model, which is better suited for modeling *skill acquisition*.

We then optimize the profile of the practice process for four specific variants of the model, depending on (i) whether the effect of practice on performance is additive or multiplicative, (ii) whether the base level adapts to past practices as the geometric mean of past intensities or as their maximum. For each variant, we consider two scenarios, namely whether the effort associated with the ultimate practice period (i.e., the final performance test) is required to be high or not.

We find that, for the additive model of performance (skill acquisition), the optimal process involves *phases* of successive increases and decreases in intensity, consistent with the principle of distributing practice to enhance learning (Brown et al. 2014), yielding U-shaped effort. Also, the transitions in practice intensity are smoother for skills involving short-term memory (i.e., geometric adaptation) as for cognitive skills, than for those involving long-term memory (i.e., maximum adaptation) as for continuous motor skills. In the latter case, the optimal process prescribes a pronounced period of rest before the ultimate performance test if it is demanding, akin to the "tapering" principle in endurance sports.

For the multiplicative model of performance (skill retention), the optimal practice process consists of *cycles* of either constant effort or pulsed effort up to the final performance test, demonstrating the power of habit. Constant effort is optimal for skills that involve short-term memory (i.e., geometric adaptation), as is the case for cognitive skills. Pulsed effort is optimal before a challenging final performance test for skills that involve long-term memory (maximum adaptation), as is the case for continuous motor skills, consistent with the principle of alternating stress and rest (Stulberg and Magness 2017) and periodization in sports.

The paper is organized as follows. We position our contribution relative to the extant literature in §2 and propose a model of practice process in §3. In §4, we optimize the practice process model to maximize performance on a given date and characterize the structure of its optimal solution. We present our conclusions in §5. Appendices EC.1 and EC.2 calibrate the model with running data and Appendix EC.3 contains the proofs and supporting lemmas.

2. Literature review

This paper relates to the literatures on endurance sports training, learning, and people-centric operations management. We build on the first two streams by bringing a dynamic optimization perspective to training and learning, and contribute to the third stream by introducing a novel domain for process analysis and optimization.

2.1. Training for Endurance Sports

Although operations research has been applied to sports optimization (see Machol et al. 1976 and Gerchak 1994 for reviews), much of its focus is on in-game strategies, drafting, and team scheduling. In contrast, training optimization has received less attention; Ladany (1975) and Zwols and Sierksma (2009) are two notable exceptions, but they consider static training models.

In contrast, a large body of literature in physiology has modeled the dynamic effects of practice on performance. In a series of papers, Banister et al. (1975) and Calvert et al. (1976) propose an analytical model that related training intensity to performance through a positive effect ("fitness") and a negative effect ("fatigue"). Adapting their model to a discrete-time setting, Morton et al. (1990) and Fitze-Clarke et al. (1991) showed that practice sessions do not all contribute equally to performance. In particular, they derived analytical formulas when practice makes the biggest contribution to performance, and when reducing practice intensity (or "tapering") enhances performance. This model is taken as a starting point for our study.

The fitness-fatigue model has been used to predict athletes' performance in a variety of endurance sports, including swimming (Banister et al. 1975, Calvert et al. 1976, Hellard et al. 2005), running (Morton et al. 1990, Wallace et al. 2014, Wood et al. 2005), weight-lifting (Busso et al. 1990), triathlon (Banister et al. 1999), and hammer-throwing (Busso et al. 1994), among others. To improve its descriptive ability, various enhancements have been proposed such as nonlinearities (Hellard et al. 2005) or multiplicative effects of fitness and fatigue on performance (Moxnes and Hausken 2008), which are also incorporated in our model. For more details on the fitness-fatigue model and its descriptive ability, see Smith (2003) and Taha and Thomas (2003).

In addition to developing mathematical models, the physiology literature provides ample evidence of high-performance strategies such as "tapering," i.e., progressively reducing the training load before a race (Mujika and Padilla 2003), "periodization," i.e., structuring the training program in cycles of high-intensity sessions followed by periods of recovery (Bompa and Haff 2009), and lifelong exercise for long-term benefits (Gries et al. 2018). However, few attempts have been made to demonstrate the optimality of such strategies besides Morton (1991) and Banister et al. (1999), who evaluated the performance of a specific set of training profiles via simulation.

Our paper therefore bridges the gap between the operations research literature on (static) training optimization and the physiology literature on the dynamic effects of practice on performance, using a dynamic model of the practice process to maximize performance.

2.2. Learning

Because learning and memory is one of the central topics in psychology, we summarize some of the key results below. Terry (2009) and Anderson (2000) both offer comprehensive reviews and Brown et al. (2014) provides a summary targeted to a broader audience. Schmidt and Lee (2011) review the state of the art in motor learning, which has a strong connection with cognitive learning (Schmidt and Bjork 1992). See also Staats (2018) for a behavioral treatment of the topic.

Combining outcomes from experiments across various settings (motor skills, memory, complex tasks), Newell and Rosenbloom (1981) formulate the "power law of practice" according to which performance improves with the number of trials as a power function. Building on Chambliss (1989), Ericsson et al. (1993) find a correlation between musicians' expertise and their cumulative practice and suggest that expertise in contexts as varied as music, chess, typing, sports, and science is driven less by innate ability than by the total amount of deliberate practice.

A second fundamental result is the "power law of forgetting" according to which a memory decays as a power function of the time it is retained (Wixted and Carpenter 2007), although some early work proposed that memory decays exponentially (Ebbinghaus 1913, Heathcote et al. 2000). In this paper, we consider exponential decay as an approximation to power decay.

In operations management, the learning curve predicates a reduction in cost as volume of production accumulates (Wright 1936), with significant implications for operations strategy (Henderson 1984, Hax and Majluf 1982). However, forgetting may happen during interruptions; see Jaber and Bonney (1997) for a review of learning-forgetting models.

The learning literature often makes the distinction between performance at the end of the skill acquisition stage and long-term retention (Schmidt and Bjork 1992). In particular, some behavioral responses to training may be only temporary. In particular, Schmidt and Bjork (1992, p. 207) argue that "typical training procedures are far from optimal" in both the motor and verbal domains because they tend to focus on short-term performance as opposed to long-term learning. As a result, these authors recommend measuring performance during a *retention* (or *transfer*) period that occurs after an extended period of rest during which no stimulus is applied, so as to measure the long-term effects of training. Although we do not formally impose a rest period in our model, practice intensity could easily be constrained to zero for a period of time before the final performance assessment. Overall, the work by Schmidt and Bjork (1992) motivates taking an optimization perspective on practice processes, applicable across motor and cognitive domains, which is the focus of this paper.

To counteract forgetting, multiple repetitions are usually necessary. Research suggests that spacing repetitions, i.e., distributing practice, is more conducive to long-term retention than concentrated, or massed, practice (Bourne and Archer 1956), although it may lead to lower performance in the short term (Schmidt and Bjork 1992); see reviews by Crowder (1976), Cepeda et al. (2006), and Brown et al. (2014). This may be because resting shifts the forgetting curve, a phenomenon originally called "reminiscence" (Williams 1926, Eysenck 1965), but now attributed to memory consolidation (Brown et al. 2014), leading Cuddy and Jacoby (1982) to observe that "forgetting improves remembering." However, it could also be because the retrieval practice following a period of inactivity is more effortful (Roediger and Karpicke 2006). Our model captures both effects.

To explain these phenomena, various theories of learning and memory have been proposed from both behavioral and cognitive perspectives (Anderson 2000). We take a behavioral perspective, in the same spirit as Hull (1943), whose theory continues to influence the development of modern theories of learning and memory, despite its limitations (Anderson 2000). In particular, our model relates to the dual-process theory of learning (Groves and Thompson 1970) since our dual constructs of *adaptation* and *stock increase* parallel their constructs of *habituation*, which happens in the stimulus-response pathway, and *sensitization*, which affects the state of the system. Our model also relates to the cognitive theory of disuse (Bjork and Bjork 1992) in the sense that our dual concepts of fitness and fatigue (stocks) and base level (regulating flow) echo their concepts of storage (stock) and memory retrieval (regulating flow).

Most of the literature on learning and memory is scientific in the sense that is phenomenological, with the exception of Newell and Simon (1972), which relates more to artificial intelligence and engineering, with limited attention given to optimization of learning besides Atkinson (1972) who, building on Markov models of learning (Bernbach 1965), proposes an optimization model for sequencing words to study. Our contribution to the learning literature—bringing an optimization perspective—fills a gap in the sense that there is limited guidance on the optimal amount of spacing or optimal intensity of each practice. We hope our analytical model will stimulate new practice process configurations to be tested experimentally.

2.3. People-Centric Operations Management

Although research in operations management has traditionally considered people as fixed entities, a growing stream has investigated how people interact with processes and its impact on operational outcomes, in the tradition of time-and-motion studies by Taylor (1911) and work-curve studies (Kraepelin 1902, Thorndike 1912), thanks in part to the greater availability of granular data sets.

Below we review the empirical literature on the effects of fatigue and adaptation on employee productivity. While our model accounts for both effects, our focus is on short-term performance and not long-term productivity.

On fatigue, KC and Terwiesch (2009), Powell et al. (2012), Green et al. (2013), Tan and Netessine (2014), Dai et al. (2015) and Xu et al. (2017) investigate the effects of overload and shift duration on employee performance, absenteeism, and process compliance. In particular, Dai et al. (2015) report a decrease in process compliance between the beginning and the end of a shift and higher compliance rates after longer breaks, which they attribute to fatigue.

On adaptation, Staats and Gino (2012) and Ibanez et al. (2017) study the impact of task specialization and task batching on productivity. In particular, Staats and Gino (2012) report that, within a day, productivity increases when workers specialize. Similarly in our model, repeated practice leads to adaptation, hence requires less subsequent effort.

In terms of modeling approach, ours is closely related to the service experience model of Das Gupta et al. (2015) in the sense that we model human response to time-varying stimuli (i.e., practice) and formulate the resulting process optimization problem as a fixed-horizon dynamic optimization problem. Our model of behavioral response includes elements of memory decay (Nerlove and Arrow 1962, Das Gupta et al. 2015), satiation (Baucells and Sarin 2007, 2010), or more specifically fatigue (Baucells and Zhao 2018), and adaptation (Wathieu 1997, Popescu and Wu 2007, Baucells and Sarin 2010, Nasiry and Popescu 2011, Aflaki and Popescu 2013, Das Gupta et al. 2015). Although most adaptation models take as a reference point the arithmetic mean of past stimuli, with the exception of Nasiry and Popescu (2011) who consider the minimum stimulus experienced so far, we consider a geometric mean and a maximum mean. Obviously, other behavioral effects could be incorporated into the model; see (Karmarkar and Karmarkar 2014) for an overview of common behavioral responses. For a specific variant of our model, it may be optimal to alternate high- and low-intensity practices, similar to advertising pulsing (Simon 1982) or service delivery with gain-seeking consumers (Aflaki and Popescu 2013).

To the best of our knowledge, practice has not been investigated from a process optimization perspective. We thus contribute to the literature on people-centric operations by developing a novel process model for optimizing learning and performance, incorporating the well-documented effects of fatigue and adaptation.

3. Model

We first present the fitness-fatigue model proposed by Banister et al. (1975) in the context of endurance sports and discuss its relevance to motor and cognitive learning. We argue that the model is incomplete, in the sense that, when subject to optimization, it leads to unrealistic solutions. We thus propose several modifications to make it more amenable to process optimization.

3.1. Fitness-Fatigue Model

Consider a practice process consisting of T evenly spaced impulse practice sessions with varying intensities w_t, $t = 1, \dots, T$, assumed to be unidimensional.[2] Throughout this paper, we denote the entire practice process as a vector $\mathbf{w} \doteq (w_1, \dots, w_T)$ or as a sequence $\{w_t\} \doteq \{w_1, \dots, w_T\}$.

Practice has two opposite effects on performance: On the one hand, practice increases strength in fitness; on the other hand, practice develops fatigue. Thus, the flows of practice contribute to two stock variables, namely *fitness* F_t and *fatigue* G_t. Without practice, these stocks are assumed to decay exponentially, either due to forgetting (for fitness) or recovery (for fatigue). Denoting by α and β the decay rates of fitness and fatigue, with $\alpha, \beta \in (0, 1)$, we thus have, for any $t = 1, \dots, T$,

$$F_t = \alpha F_{t-1} + w_t \tag{1}$$

$$G_t = \beta G_{t-1} + w_t, \tag{2}$$

for some $F_0, G_0 \geq 0$.

Performance increases with fitness and decreases with fatigue. Considering a linear relationship, Banister et al. (1975) formulate performance in period t (P_t) as:

$$P_t = P_0 + k_F F_t - k_G G_t, \tag{3}$$

where $k_F, k_G \geq 0$. As reviewed in §2.1, the model has descriptive ability for a variety of endurance sports. Moreover, the parameters' estimated values have proven to be robust across numerous

[2] In endurance sports, the intensity of a particular session is usually measured as the product of its stress (duration, distance) and its strain (measured on a cardinal scale, e.g., cardiovascular intensity, or on an ordinal scale based on the athlete's input). In cognitive learning, intensity can be measured on an ordinal scale taking into account the following: (i) the intrinsic difficulty of the material (e.g., single-digit multiplications are simpler than multiple-digit ones); (ii) the duration of the session; and (iii) the level of participation (passive encoding, e.g., listening, has lower intensity than active encoding, e.g., studying, which has itself lower intensity than retrieval, e.g., being tested, explaining; see Terry 2009, Brown et al. 2014). Motor learning would lie between these two extremes.

studies (Taha and Thomas 2003). For these reasons, it has been implemented in software to provide performance feedback to athletes during training (Coggan 2008).

In this model, not only does the cumulative amount of practice matter but also its *timing*. Indeed, poorly-timed practice may have a detrimental effect on performance. Specifically when $\alpha > \beta$, practice in the last periods of the practice process, namely, when $t > T - \ln(k_G/k_F)/\ln(\alpha/\beta)$, has a negative effect on performance P_T because $\partial P_T(\mathbf{w})/\partial w_t < 0$ in those periods, demonstrating the optimality of *tapering*, i.e., reducing the intensity of practice in the final days before a race.

3.1.1. Applicability to Motor and Cognitive Learning. Hull (1943) proposes a conceptual behavioral model of learning to describe an individual's "effective reaction potential," i.e., the probability, speed, or force with which an individual's behavior would be performed in response to a stimulus. In Hull's model, repeated practice of a stimulus-response association strengthens a *habit*, which, like fitness, increases the reaction potential, but also develops an *inhibitory potential*, which like fatigue, diminishes the reaction potential. Similar to (1)-(2), both the habit strength and the inhibitory potential increase with the number of repetitions, and decay exponentially with the passage of time (Hull 1943). At a high level, the model is consistent with the popular stress and rest framework for peak performance in sports, science, and art (Stulberg and Magness 2017).

Model (1)-(3) captures many salient features of cognitive learning processes. First, according to the *power law of learning*, memory strength increases as a power (or exponential, see Heathcote et al. 2000) function of practice (Anderson 2000). With constant intensity $w_t = w$ and no fatigue effects ($k_G = 0$), we obtain $P_t = P_0 + k_F w(1-\alpha^t)/(1-\alpha)$, which increases exponentially in the number of trials. Second, according to the *power law of forgetting*, memories decay as a power (or exponential, see Ebbinghaus 1913) function of the time over which they are being retained (Anderson 2000). With $w_t = 0$ for $t > 1$ and no fatigue effects ($k_G = 0$), we indeed obtain that $P_t = k_F w_1 \alpha^{t-1}$, which decays exponentially in time. Third, recall capability may improve after some rest, because of memory consolidation (Brown et al. 2014). Similarly research on conditioning has reported phenomena of recall, reinstatement, and spontaneous recovery (Bjork and Bjork 1992). Such phenomena can be described with Model (1)-(3) as an improvement of performance after a period of rest due to the dissipation of the inhibitory potential (or fatigue) during rest (Eysenck 1965). Fourth, *distributed* practice is often more effective than *massed* practice (Schmidt and Lee 2011, Brown et al. 2014). Using (1)-(3), one can construct a problem instance where spaced practice leads to greater performance than massed practice both at the end of the acquisition period *and* after an extended period of rest, during the retention period (cf. discussion in §2.2).[3]

[3] Consider a setup similar to Bourne and Archer (1956) for motor learning with 21 30-second trials, either performed consecutively or interspersed by rest periods of 1 minute, then followed by 5 minutes of rest, and then by another

3.1.2. Practice Process Optimization. Although scientific experimentation has uncovered high-performing strategies (e.g., tapering, distributing practice), little is known about the optimal profile of the practice process; moreover, intuition is often a poor guide to identify what strategies may work best (Simon and Bjork 2001). For these reasons, optimizing a model with good descriptive ability, such as the fitness-fatigue model, could help refine commonly used high-performance practice strategies and identify the conditions under which they are indeed optimal.

In what follows, we assume that practices are free (e.g., there is no opportunity cost of time) and nonnegative, i.e., $w_t \geq 0$ for all $t = 1, \ldots, T$. In addition, we assume that the last practice (which could consist of a performance test like a race or an exam) has a lower bound $\underline{w}_T$ (potentially set to zero),[4] and we optimize the practice process to maximize performance in that period, P_T.[5]

Given our objective and constraints, the fitness-fatigue model (1)-(3) appears incomplete, in the sense that its optimized solution does not seem realistic. First, due to the model's linear nature, maximizing $P_T(\mathbf{w})$ may result in an unbounded solution. Imposing upper bound constraints on the practice intensity, i.e., $0 \leq w_t \leq \overline{w}_t$, would lead to the following solution when $\alpha > \beta$: $w_t = \overline{w}_t$ for all $t < T - \ln(k_F/k_G)/\ln(\alpha/\beta)$ and $w_t = 0$ for all $t > T - \ln(k_G/k_F)/\ln(\alpha/\beta)$. That is, the optimal policy is a threshold policy. In this case, the optimal solution would be highly dependent on the choice of upper bounds $\overline{w}_t$, which may be arbitrary. Second, the model would recommend exerting no effort in the last few sessions, which few athletes would find credible. Third, the model would not tailor the practice process to the type of performance assessment. Fourth, the threshold nature of the optimal policy starkly contrasts with the principles of distributing practice or periodization, despite strong empirical support. In the next section, we enhance the model to alleviate some of these shortcomings.

3.2. A Generalized Fitness-Fatigue Model

We modify the fitness-fatigue model in three ways: First, we incorporate the effect of adaptation to practice. Second, we introduce nonlinearities to capture diminishing marginal returns of practice

9 consecutive trials. Define a period as a 30-second interval and consider the following parameters (so as to match the reported data): $P_0 = 36.785$, $k_F = 3.0489$, $\alpha = 1$, $k_G = 12.902$, $\beta = 0.8286$ and $w_t = 1$ during a trial and $w_t = 0$ otherwise. Under this calibrated model (1)-(3), performance in trials 21 and 22-30 is uniformly the highest under the 1-minute rest condition and uniformly the lowest under the no-rest condition, consistent with their reported results.

[4] Similar to Footnote 2, the last-period intensity requirement $\underline{w}_T$ can be defined as a combination of stress and strain.

[5] Our approach can be extended to other objectives or incorporate additional constraints. For instance, there could be a cost or upper bound on practice intensity in the event that time is limited. In addition, there could be an upper bound on fatigue, e.g., to prevent injury. Third, there could be intermediate performance objectives (e.g., mid-term exams), in which case the objective should take the weighted average of performance at difference time epochs. Finally, research on learning has shown that strategies that maximize performance at the end of the acquisition period may differ from strategies that maximize learning, i.e., performance after a retention period (Schmidt and Bjork 1992). To focus on long-term learning, we could impose that $w_t = 0$ for $t = \hat{t}, \ldots, T-1$, for some $\hat{t}$, so that the temporary effects of learning have time to dissipate. We leave it for future research to explore such extensions.

on fitness and increasing marginal costs on fatigue. Third, we propose that the effects of fitness and fatigue can be either additive or multiplicative. We outline these modifications separately and then combine them in a model that we calibrate in Appendix EC.1.

3.2.1. Adaptation. Although model (1)-(3) considers a uniform effect of practice on performance, there is ample evidence of adaptation to repeated practice (Groves and Thompson 1970). In physical training, "the benefit of [a] new training stress gets less as time goes by ... and eventually, if you just continue doing the same training week after week, your new fitness level will no longer improve" (Daniels 2014, p. 18). Similarly, "as people learn a motor skill, they appear to do the task with less and less physical and mental effort" (Schmidt and Lee 2011, p. 340). Likewise for cognitive skills (e.g., solving geometry problems), "as students become more practiced in a skill, they come to recognize directly what they formerly had to think through" (Anderson 2000, p. 320).

In the spirit of adaptation models (Wathieu 1997) and consistent with the concept of deliberate (or effortful) practice (Ericsson et al. 1993), we measure the impact of w_t on fitness F_t and fatigue G_t relative to a *base level*, denoted as b_t. Specifically, we define the ratio w_t/b_t as *effort*. See Morton et al. (1990) and Wallace et al. (2014) for dynamic normalizations of the effects of practice intensity due to subjects' changing physiology and Busso et al. (1997) for a dynamic model recalibration. When b_t is constant for all t, this adaptation model reduces to the original fitness-fatigue model.

We assume that the base level is the mean intensity of past practices. Thus, less effort is required if one has practiced at high intensity (e.g., long runs, complex geometry problems) in the past, especially in the recent past. Specifically,

$$b_t = \epsilon \cdot h\left(w_{t-1}, b_{t-1}\right), \tag{4}$$

for $t > 1$ with $b_1 > 0$, in which $h(w,b)$ is a generalized mean (Hardy et al. 1952), and $0 < \epsilon \leq 1$ measures some decay in the base level.

For analytical tractability, we consider the *geometric mean*, i.e., $h(w,b) = w^\theta b^{1-\theta}$ with $0 \leq \theta \leq 1$, and the *maximum mean*, i.e., $h(w,b) = \max\{w,b\}$. Adaptation to high- (low-)intensity practices will be faster (slower) under the maximum mean than under the geometric mean. Given that verbal-cognitive components tend to be more quickly forgotten than motor components (Schmidt and Lee 2011) and that, within motor skills, discrete tasks (e.g., throwing, shifting gears in a car) tend to be more quickly forgotten than continuous tasks (e.g., swimming, riding a bicycle), the geometric mean model seems more relevant for cognitive tasks and the maximum mean model seems more relevant for continuous motor tasks, with discrete motor tasks in between. Of course, this is a generic guideline; the choice of the most appropriate model must be context-dependent.[6]

[6] Alternatively, one could categorize skills along a continuous spectrum. In that case, a generalized mean function (e.g., $h(w,b) = (\theta w^r + (1-\theta)b^r)^{1/r}$ for $r \geq 0$) could map skills to specific values of the elasticity parameter (r). Moreover, adaptation could happen in discrete steps (e.g., capability milestones or stratification, see Chambliss 1989) in certain contexts. We leave it for future research to characterize the optimal processes under such alternative models.

3.2.2. Nonlinearities. The extreme nature of the optimal solution of the optimized fitness-fatigue model (1)-(3) comes from its linearity, yet there is ample evidence of nonlinearities in the response of fitness and fatigue to effort (Hellard et al. 2005). In particular, Daniels (2014) formulates principles of training intensity's diminishing returns on fitness but accelerating setbacks, suggesting concave effects on fitness and convex effects on fatigue. In cognitive processes, satiation or mental fatigue within a practice session could lead to similar nonlinear effects.

For analytical tractability, we consider here power functions. Specifically, ignoring the effects of adaptation, the effect of a practice session with intensity w_t on fitness F_t is equal to w_t^{λ} for some $\lambda < 1$ and its effect on fatigue G_t is equal to w_t^{μ} for some $\mu > 1$.

3.2.3. Additive or Multiplicative Impact of Practice. In contrast to the original model (1)-(2), which assumes an additive effect of practice on fitness and fatigue, there may be settings where the effect is multiplicative. In the spirit of the principle of accelerating setbacks (Daniels 2014), the higher the fatigue, the more detrimental the impact of an intensive workout. Similarly in cognitive processes, retrieval practices have a compounding effect on past memories through a "reconsolidation" process, which "helps to reinforce meaning, strengthen connections to prior knowledge, bolster the cues and retrieval routes for recalling it later, and weaken competing routes" (Brown et al. 2014, p. 97). In fact, activation theories propose that retrieving a subset of some studied material should facilitate retrieval of the remaining material, provided that some associative links exist between them (Chan et al. 2006). Accounting for such effects, a multiplicative version of the fitness-fatigue model could then be expressed as:

$$F_t = F_{t-1} \cdot (\alpha + w_t) \tag{5}$$

$$G_t = G_{t-1} \cdot (\beta + w_t) \, . \tag{6}$$

In addition, instead of defining performance as the difference between fitness and fatigue, as in (3), one could define performance as a function of their ratio:

$$P_t = P_0 + \gamma \cdot \frac{F_t}{G_t}, \tag{7}$$

where $\gamma > 0$; see Moxnes and Hausken (2008) and Wallace et al. (2014) for similar models.

Because early skill acquisition is needed before practice can have a compounding effect, the additive and the multiplicative models capture different stages of skill development. Unlike the additive model (1)-(2), which can be initialized with $F_0 = G_0 = 0$ and gradually builds up through practice, the multiplicative model (5)-(6) presumes that $F_0 > 0$ and $G_0 > 0$ so as to lead to compounding. Thus, the additive model may be better suited to the early stages of skill development, i.e., the skill *acquisition* stage, in which new material is added to the stocks of fitness (or memory strength) and fatigue; whereas, the multiplicative model may be better suited to later stages, i.e., the skill *retention* stage through on-the-job experience development or subsequent training.

3.2.4. General Formulation. Combining the effects of adaptation and nonlinearities into the additive model (1)-(3) and the multiplicative model (5)-(7), we obtain:

$$P_t = P_0 + k_F F_t - k_G G_t, \qquad F_t = \alpha F_{t-1} + \left(\frac{w_t}{b_t}\right)^{\lambda}, \qquad G_t = \beta G_{t-1} + \left(\frac{w_t}{b_t}\right)^{\mu} \tag{8}$$

$$\text{or } P_t = P_0 + \gamma \frac{F_t}{G_t}, \qquad F_t = F_{t-1} \cdot \left(\alpha + \left(\frac{w_t}{b_t}\right)^{\lambda}\right), \qquad G_t = G_{t-1} \cdot \left(\beta + \left(\frac{w_t}{b_t}\right)^{\mu}\right), \tag{9}$$

in which $b_t = \epsilon \cdot h(w_{t-1}, b_{t-1})$ and either $h(w,b) = w^{\theta} b^{1-\theta}$ or $h(w,b) = \max\{w,b\}$. Hence, we focus on the following four variants of the model with the following interpretations:

- Additive model with geometric mean: Acquisition of cognitive skills,
- Additive model with maximum mean: Acquisition of continuous motor skills,
- Multiplicative model with geometric mean: Retention of cognitive skills,
- Multiplicative model with maximum mean: Retention of continuous motor skills.

In Appendix EC.1, we illustrate how the model can be calibrated with readily-available data from a runner's training program. For this particular data set, we find that all four variants of the model have similar goodness of fit. Moreover, adaptation plays a more important role than the differential in decay parameters (α and β) and nonlinearities (λ and μ). Even though the proposed enhancements have a marginal impact on the fitting capability of the model (on this data set), their primary goal remains to make it more amenable to optimization, as we explore next.

4. Optimal Practice Process Profiles

Using the four variants of the model introduced in the last section, we next optimize the practice process $\mathbf{w}$ to maximize performance in period T, $P_T(\mathbf{w})$. We assume that the intensity of the last practice (e.g., an exam, a race) must have at least intensity $\underline{w}_T$ and require that the intensity of the preceding practice sessions be nonnegative. We first consider the additive model (8) and then the multiplicative model (9).

4.1. Additive Model (Skill Acquisition)

We first consider the additive model of performance (8), which, as discussed in §3.2.3, fits well the skill acquisition stage of skill development. For the purpose of optimizing the profile of the practice process, we assume without loss of generality that $P_0 = 0$, $k_G = 1$, and $k_F = \gamma > 0$. Accordingly, the optimization problem for the additive model ('+') can be expressed as follows:

$$
\begin{aligned}
\underset{\mathbf{w},\mathbf{b},\mathbf{F},\mathbf{G}}{\text{maximize}} \quad & \gamma F_T - G_T \\
\text{subject to} \quad & F_t = \alpha F_{t-1} + \left(\frac{w_t}{b_t}\right)^{\lambda} \quad && \forall t > 1, \\
& G_t = \beta G_{t-1} + \left(\frac{w_t}{b_t}\right)^{\mu} \quad && \forall t > 1, \\
& b_t = \epsilon \cdot h(w_{t-1}, b_{t-1}) \quad && \forall t > 1, \\
& w_T \geq \underline{w}_T, \\
& w_t \geq 0 \quad && \forall t \geq 1.
\end{aligned} \tag{P_+}
$$

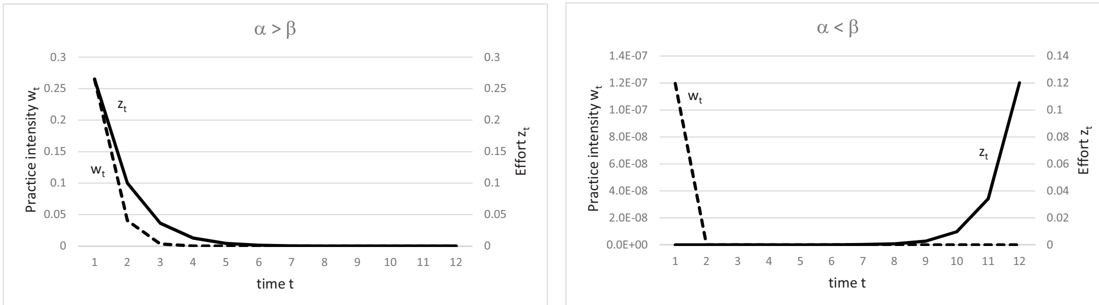
We define the *(relative) effort* z_t as the ratio of the current practice intensity w_t over the mean intensity of the past practice sessions b_t, i.e., $z_t \doteq w_t/b_t$. With that change of variable, after substituting the definitions of F_t and G_t into the objective, the problem can be simplified as follows:

$$
\begin{aligned}
\underset{\mathbf{z}}{\text{maximize}} \quad & \gamma \alpha^T F_0 - \beta^T G_0 + \sum_{t=1}^{T} \left(\gamma \alpha^{T-t} z_t^{\lambda} - \beta^{T-t} z_t^{\mu}\right) \\
\text{subject to} \quad & z_T \prod_{t=1}^{T-1} h\left(z_t, 1\right) \geq \frac{\underline{w}_T}{\epsilon^{T-1} b_1}, \\
& z_t \geq 0 \qquad \forall t > 1.
\end{aligned} \tag{P'_+}
$$

The structure of the optimal solution depends on whether the coupling constraint in (P'_+) is loose or tight at the optimum. For given effort $\mathbf{z}$, the constraint will tend to be loose if the learner's initial abilities (b_1) are high relative to the intensity of final performance assessment ($\underline{w}_T$), and if they do not decay too quickly (i.e., if ϵ is close to 1). Given that $\max\{w, b\} \geq w^{\theta} b^{1-\theta}$, for any $\theta \in [0, 1]$, the constraint will also tend to be looser for skills that have strong memory (e.g., continuous motor skills) than for those that have short memory (e.g., cognitive skills). In either case, we will show that the optimal practice process for skill acquisition is structured in *phases* of practice increase and decrease, consistent with the principle of distributing practice (Brown et al. 2014).

4.1.1. Low Final Effort Requirement. We first consider the case of a low final effort requirement, i.e., when the coupling constraint in (P'_+) is loose at the optimum. In that case, problem (P'_+) is additively separable; thus, efforts are independent of each other. In particular, deviating from the optimal solution in one period (i.e., exerting more or less effort than optimal) will not require adjusting the following periods' efforts. Examining the objective function in (P'_+) reveals that the fundamental trade-off in determining $\mathbf{z}$ lies in the relative values of the fitness and fatigue decay rates α and β, similar to the original model (1)-(3). The next proposition characterizes the optimal profile of the practice process and Figure 1 illustrates the optimal effort $\{z_t^*\}$ (solid lines) and practice intensity $\{w_t^*\}$ (dashed lines) both when $\alpha \geq \beta$ and when $\alpha < \beta$. Condition (C_+) below is obtained by first solving (P'_+) without the coupling constraint, and then plugging that solution into the coupling constraint.

Figure 1 Optimal Effort and Practice Intensity under the Additive Model with Low Final Effort Requirement



Note. $h(w,b) = w^{\theta} b^{1-\theta}$, $\alpha = 0.7$, $\gamma = 0.8$, $\theta = 0.6$, $\epsilon = 0.9$, $\underline{w}_T = 0$, $b_1 = 1$, $\lambda = 0.9$, $\mu = 1.1$, $T = 12$, $\beta = 0.69$ (left) and $\beta = 0.9$ (right).

PROPOSITION 1. *For Problem (P_+), when the ultimate effort requirement is low, i.e., when the following condition holds:*

$$\left(\frac{\gamma\lambda}{\mu}\right)^{\frac{1}{\mu-\lambda}} \prod_{t=1}^{T-1} h\left(\left(\frac{\gamma\lambda\alpha^{T-t}}{\mu\beta^{T-t}}\right)^{\frac{1}{\mu-\lambda}}, 1\right) \geq \frac{\underline{w}_T}{\epsilon^{T-1} b_1}, \tag{C_+}$$

- *if $\alpha \geq \beta$, the optimal effort $\{z_t^*\}$ is decreasing and the optimal practice intensity $\{w_t^*\}$ is quasi-concave;*
- *if $\alpha < \beta$, the optimal effort $\{z_t^*\}$ is increasing and the optimal practice intensity $\{w_t^*\}$ is quasi-convex.*

The optimal solution, when expressed in terms of effort $\mathbf{z}$, is similar to the optimal solution of the original model (1)-(3). Specifically if $\alpha > \beta$, effort $\{z_t\}$ must be decreasing, consistent with the often recommended tapering strategy; whereas if $\alpha < \beta$, effort should be increasing since the effect of practice on fitness is short-lived compared to its effect on fatigue. However, due to the introduced non-linearities, the effort transitions are smoother than in the bang-bang solution of the linear model discussed in §3.1.2.

The dynamics of practice intensity $\{w_t\}$ reflect the two ways practice affects performance. On the one hand, practice at time t has a *direct* effect on both fitness and fatigue at time t, and their combined effect on P_T depends on their respective decay rates. On the other hand, it has an *indirect* effect on P_T through a change in the base level b_t, which, if it increases, makes future practice more effective at preventing fatigue, but less effective at further developing fitness. Thus, a high-intensity practice session has a positive effect on performance because it increases current fitness and makes future practice less prone to fatigue, but also a detrimental effect because it increases fatigue and makes future practice less effective at further increasing fitness.

When fitness decays more slowly than fatigue, i.e., $\alpha \geq \beta$, the optimal profile of the practice process $\{w_t\}$ is quasi-concave. Hence early in the process, effort $\{z_t\}$ is decreasing despite an increase in practice intensity $\{w_t\}$. These different dynamics arise because of adaptation: Increasing practice intensity develops the base level b_t, which makes subsequent practices less effortful. Towards the end of the process, effort and practice intensity are both decreasing and thus aligned.

To illustrate this profile of the practice process, consider a student learning how to multiply. When first exposed to multiplications, this student may initially struggle with the novelty of the topic, i.e., z_t is large, even though the material (e.g., single-digit multiplications) is technically simple. As time progresses, new material (e.g., multiple-digit multiplications) is introduced and the student may be required to participate more actively (e.g., solving problems), i.e., $\{w_t\}$ is initially increasing. However, as she accumulates practice, the student will make fewer and smaller errors and gradually switch from a cognitive stage (slow and deliberate use of knowledge) to an associative stage (more direct representation of what to do) (Anderson 2000); that is, $\{z_t\}$ is decreasing. At some point, no new material is added and the retrieval sessions may be shorter, i.e., $\{w_t\}$ starts decreasing, and the student's skills become more automated and rapid, i.e., $\{z_t\}$ keeps decreasing.

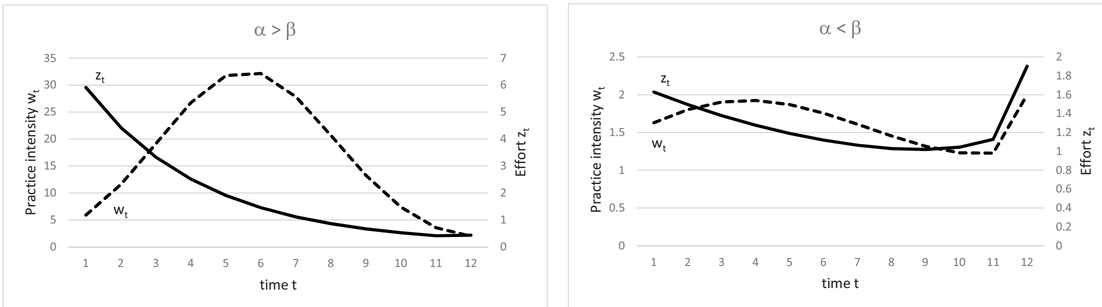
In contrast, when $\alpha < \beta$, fitness decays faster than fatigue, and the optimal profile of the practice process $\{w_t\}$ is quasi-convex. Early in the process, effort thus increases despite a decrease in process intensity. This different dynamics can again be explained by adaptation: decreasing practice intensity lowers the base level b_t, which makes subsequent practices more effortful. Towards the end of the process, effort and practice intensity are both increasing and thus aligned. To understand why it may be optimal to have the base level decrease initially, recall that memories are short-lived when $\alpha < \beta$ (e.g., sensory memories; see Anderson 2000). One thus needs to give the last set of practices as much attention as possible for these memories to be effectively captured. The decrease in practice intensity characterizing the early part of the process thus aims to reset the base level at a low value to build up the amount of future attention available and capture future memories more effectively, similar to a wine taster who may need to reset her palate to give a wine full attention.

In sum, these two cases show that even though effort should be monotone, the intensity of practice may not be so, and it is in general unimodal. In particular reducing practice intensity may be optimal to either let fatigue recover while preserving most fitness (when $\alpha > \beta$) or let the base level reset for maximizing the benefit of future practice (when $\alpha < \beta$).

4.1.2. High Final Effort Requirement. We next consider the case of a high final effort requirement, i.e., when the coupling constraint in (P'_+) is tight at the optimum. In that case, effort $\mathbf{z}$ should be guided not only by the difference in decay rates between fitness and fatigue, but also by the preparation of the base level b_t for the ultimate effort requirement.

We consider separately the cases of geometric and maximum adaptation. Overall, we find that effort should in general be U-shaped. Hence, the high ultimate effort requirement introduces two

Figure 2 Optimal Effort and Practice Intensity under the Additive Model with Geometric Adaptation and High Final Effort Requirement



Note. $\alpha = 0.7$, $\gamma = 0.8$, $\theta = 0.6$, $\epsilon = 0.9$, $\underline{w}_T = 2$, $b_1 = 1$, $\lambda = 0.9$, $\mu = 1.1$, $T = 12$, $\beta = 0.69$ (left) and $\beta = 0.9$ (right).

types of distortion in the (monotone) effort patterns characterized in Proposition 1, namely an increase in effort at the end of the process (when $\alpha \geq \beta$) and a decrease in effort at the beginning of the process (when $\alpha < \beta$).

Geometric Adaptation (Cognitive Tasks). We first consider the case with geometric adaptation, which, as discussed in §3.2.1, tends to fit better cognitive or some discrete motor tasks. The next proposition characterizes the optimal practice process and Figure 2 illustrates the optimal effort $\{z_t^*\}$ (solid lines) and practice intensity $\{w_t^*\}$ (dashed lines) both when $\alpha \geq \beta$ and when $\alpha < \beta$.

Proposition 2. *For Problem (P_+) with $h(w,b) = w^\theta b^{1-\theta}$, when the ultimate effort requirement is high, i.e., when Condition (C_+) does not hold,*

- *if $\alpha \geq \beta$, the optimal effort $\{z_t^*\}_{t=1,\ldots,T-1}$ is decreasing (up to period T) and there exist time thresholds $t_1 \leq t_2$ such that the optimal practice intensity evolves as follows: $\{w_t^*\}_{t=1,\ldots,t_1}$ is increasing and $\{w_t^*\}_{t=t_2,\ldots,T-1}$ is decreasing.*
- *if $\alpha < \beta$, the optimal effort $\{z_t^*\}$ is quasi-convex and there exist time thresholds $t_1 \leq t_2 \leq t_3 \leq t_4$ such that the optimal practice intensity evolves as follows: $\{w_t^*\}_{t=1,\ldots,t_1}$ is increasing, $\{w_t^*\}_{t=t_2,\ldots,t_3}$ is decreasing, and $\{w_t^*\}_{t=t_4,\ldots,T}$ is increasing.*

Comparing Proposition 2 to Proposition 1 reveals that a tighter coupling constraint in (P'_+) distorts the profile of the practice process at its very end as well as in its early part. Specifically when $\alpha \geq \beta$, we find that $\{z_t^*\}_{t=1,\ldots,T-1}$ is decreasing, similar to Proposition 1, but this time, it is possible that $z_T^* > z_{T-1}^*$ due to the final performance assessment in period T. As a result, the effort profile $\{z_t^*\}$ can now be U-shaped. Similar to Proposition 1, $\{w_t^*\}$ is increasing when z_t is large, which happens in the early phase of the process, and decreasing when z_t is small, which happens in the late phase of the process, prior to period $T-1$. However, between these two phases, i.e., when z_t takes on intermediate values, the practice profile $\{w_t^*\}$ may be more convoluted. In particular, it can be shown that, in the absence of nonlinearities (i.e., $\lambda = \mu = 1$), $\{w_t^*\}_{t=1,\ldots,T-1}$

may go through up to three successive phases of decrease, increase, and then decrease. The higher effort requirement in the last period thus distorts the profile of the practice process not only at its very end, by boosting the final effort on the final performance assessment, but also in its early part, by shaping the base level in anticipation of that last challenge.

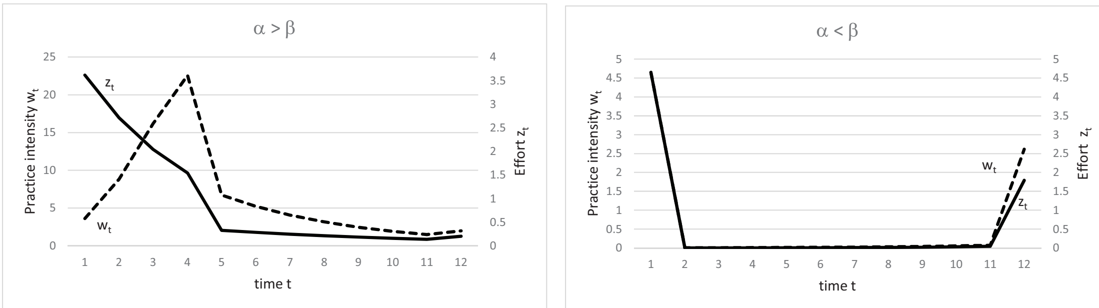
When $\alpha < \beta$, $\{z_t^*\}$ is increasing in the latter part of the practice process, similar to Proposition 1, but due to the tightening of the coupling constraint in (P'_+), it can be decreasing initially. Similar to Proposition 1, $\{w_t^*\}$ is increasing when z_t is large, which may now happen in the early and late phases of the process, and decreasing when z_t is small, which now happens towards the middle of the process. Although the profile of $\{w_t^*\}$ between these phases remains unspecified in general, it can be shown that, in the absence of nonlinearities (i.e., $\lambda = \mu = 1$), $\{w_t^*\}_{t=1,\ldots,T-1}$ may go through up to three successive phases of increase, decrease, and then increase. Hence, similar to the case where $\alpha \geq \beta$, the tightening of the ultimate effort requirement distorts the early part of the process to shape the base level in anticipation of the last-period's practice.

In sum, a higher effort requirement in the last period tends to distort not only the very end of the process by inducing higher effort in the last period (on the performance assessment), but also its early part. Moreover, the early distortion has to do less with the direct effect of practice on fitness and fatigue, because it will have mostly decayed by the end of the process, but more with its indirect effect, through the development of the base level to either make future practice more effective or less demanding. As a result, the optimal profile of the practice process $\{w_t^*\}$ may exhibit multiple phases of increase in intensity, to either build current fitness or grow the base level so as to make future fatigue more bearable; and phases of decrease in intensity, to either recover from accumulated fatigue or decrease the base level to make future practice more effective at building fitness. This pattern of practice intensity is in fact similar to the notion of spacing out (or distributing) practice to enhance learning (Brown et al. 2014), with enough time between two peaks in practice intensity to recover from fatigue (so that memories have time to consolidate) and reset the base level (so that future practice sessions are more effortful).

Maximum Adaptation (Continuous Motor Tasks). We next consider the case of maximum adaptation, which, as discussed in §3.2.1, tends to fit better continuous motor tasks. The next proposition shows that the effort pattern can be characterized by a partition, $\mathcal{S}^* = \{1,\ldots,\hat{t}\} \cup \{T\}$ for some $\hat{t}$, such that effort is high initially, i.e., $z_t^* > 1$ for all $t \in \mathcal{S}^* \setminus \{T\}$, and low subsequently up to the final performance test, i.e., $z_t^* \leq 1$ for all $t \notin \mathcal{S}^*$. Figure 3 illustrates the optimal effort $\{z_t^*\}$ (solid lines) and practice intensity $\{w_t^*\}$ (dashed lines) both when $\alpha \geq \beta$ and when $\alpha < \beta$.

Proposition 3. *For Problem (P_+) with $h(w,b) = \max\{w,b\}$, when the ultimate effort requirement is high, i.e., when Condition (C_+) does not hold, there exists an optimal partition $\mathcal{S}^* =$*

Figure 3 Optimal Effort and Practice Intensity under the Additive Model with Maximum Adaptation and High Final Effort Requirement



Note. $\alpha = 0.7$, $\gamma = 0.8$, $\epsilon = 0.9$, $\underline{w}_T = 2$, $b_1 = 1$, $\lambda = 0.9$, $\mu = 1.1$, $T = 12$, $\beta = 0.68$ (left) and $\beta = 0.9$ (right).

$\{1, \ldots, \hat{t}\} \cup \{T\}$ *for some* $\hat{t}$ *such that* $z_t^* > 1$ *for all* $t \in \mathcal{S}^* \setminus \{T\}$ *and* $0 < z_t^* \leq 1$ *for all* $t \notin \mathcal{S}^*$. *Moreover,*

- *if* $\alpha \geq \beta$, *the optimal effort* $\{z_t^*\}_{t=1,\ldots,T-1}$ *is decreasing and the optimal practice intensity* $\{w_t^*\}_{t=1,\ldots,T-1}$ *is quasi-concave;*
- *if* $\alpha < \beta$, *the optimal effort* $\{z_t^*\}_{t \in \mathcal{S}^*}$ *is quasi-convex and* $\{z_t^*\}_{t \notin \mathcal{S}^*}$ *is increasing; and the optimal practice intensity* $\{w_t^*\}_{t \in \mathcal{S}^*}$ *is structured in at most three phases: (i) increase, (ii) decrease, and (iii) increase, and* $\{w_t^*\}_{t \notin \mathcal{S}^*}$ *is monotone.*

The optimal profile of effort under maximum adaptation tends to be similar to that under geometric adaptation (Proposition 2), except that there is now a clear distinction between a phase of high-effort (in which $z_t > 1$) and a phase of low-effort (in which $z_t \leq 1$). Similar to the case of geometric adaptation (Proposition 2), a higher ultimate effort requirement distorts the optimal effort pattern (relative to the one described in Proposition 1) in two ways, namely at the very end of the process as well as in its early part. The general dynamics in the high-effort phase are similar to Proposition 2: $\{z_t^*\}$ has a U-shape and $\{w_t^*\}$ goes through at most three phases of increase, decrease, and final increase. In the particular case where $\alpha \geq \beta$, $\{z_t^*\}_{t=1,\ldots,T-1}$ is decreasing, $\{w_t^*\}_{t=1,\ldots,T-1}$ is quasi-concave, and both have a potential uptick on the final performance test.[7] Overall, the structure of the optimal solution depends less on the relative values of decays in fitness and fatigue than in the unconstrained case (Proposition 1). Therefore, the increase in the ultimate effort requirement shifts the early effects of practice on performance away from the relative accumulation of fitness and fatigue to the development of the base level.

One key difference from the case with geometric adaptation is that, under maximum adaptation, the evolution of effort $\{z_t^*\}_{t=1,\ldots,T-1}$ up to the last period T is structured into two phases, first of

[7] To illustrate this pattern, consider *Muscle & Fitness*' 60-day workout plan (`https://www.muscleandfitness.com/workouts/workout-routines/60-day-revolution-workout-plan`, last accessed on December 6, 2017) structured in six phases of modulating intensity: one week of "Intro," followed by a week of "Base," then followed by 4 weeks of "Overload," followed by a relaxing week of "Deload," and finishing with an intensive week of "Shock."

high effort (i.e., $z_t^* > 1$) and then of low effort ($z_t^* \leq 1$), and the numerical examples in Figure 3 reveal that the transitions between these two phases can be quite abrupt. In particular, there is a clear recovery phase near the end of the horizon under maximum adaptation. This is because the base level is not affected by the intensity of the low-effort sessions in that phase and only decays at rate ϵ. As a result, for skills that have long-term memory (e.g., continuous motor skills), a tapering strategy is optimal if the performance assessment is challenging.[8]

4.2. Multiplicative Model (Skill Retention)

We next consider the multiplicative model (9), which, as discussed in §3.2.3, fits well the skill retention stage. For the purpose of optimizing the profile of the practice process, we assume without loss of generality that $P_0 = 0$, $F_0 = G_0 = 1$, and $\gamma = 1$. Accordingly, the practice process optimization problem for the multiplicative model ('×') can be expressed as follows:

$$\begin{aligned}
\underset{\mathbf{w},\mathbf{b},\mathbf{F},\mathbf{G}}{\text{maximize}} \quad & F_T/G_T \\
\text{subject to} \quad & F_t = F_{t-1} \times \left(\alpha + \left(\frac{w_t}{b_t}\right)^{\lambda}\right) && \forall t > 1, \\
& G_t = G_{t-1} \times \left(\beta + \left(\frac{w_t}{b_t}\right)^{\mu}\right) && \forall t > 1, \\
& b_t = \epsilon \cdot h(w_{t-1}, b_{t-1}) && \forall t > 1, \\
& w_T \geq \underline{w}_T, \\
& w_t \geq 0 && \forall t \geq 1.
\end{aligned} \tag{$P_\times$}$$

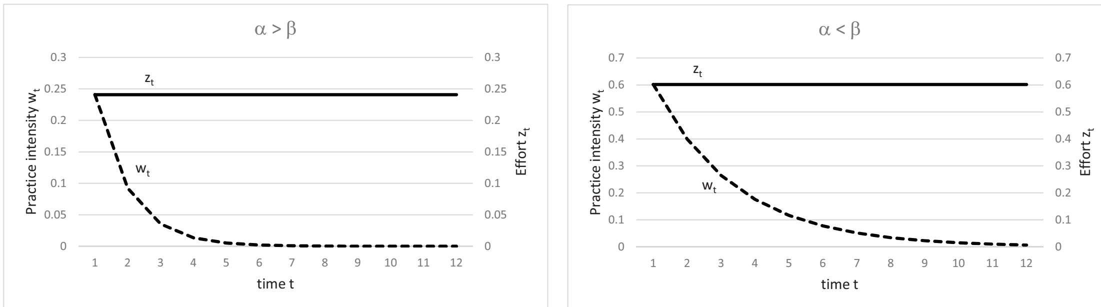
Similar to the additive model, we change variables by considering the (relative) effort $z_t \doteq w_t/b_t$. With that change of variable, after substitution of the definitions of F_t and G_t into the objective, the problem can be simplified as follows:

$$\begin{aligned}
\underset{\mathbf{z}}{\text{maximize}} \quad & \prod_{t=1}^{T} \left(\frac{\alpha + z_t^{\lambda}}{\beta + z_t^{\mu}}\right) \\
\text{subject to} \quad & z_T \prod_{t=1}^{T-1} h\left(z_t, 1\right) \geq \frac{\underline{w}_T}{\epsilon^{T-1} b_1}, \\
& z_t \geq 0 \qquad \forall t.
\end{aligned} \tag{$P'_\times$}$$

Similar to §4.1, we successively consider the cases where the coupling constraint in ($P'_\times$) is loose or tight. In both cases, we will show that the optimal practice process can be structured in *cycles*,

[8] This recovery phase towards the end of the practice process is in fact similar to the phase of no practice introduced in learning experiments to ensure that the temporary effects of learning have dissipated (Schmidt and Bjork 1992). This suggests that, under maximum adaptation (which is more relevant to continuous motor skills), a practice process that maximizes performance (which is the objective here) will be similar to one that optimizes (long-term) learning since the short-term effects are forced to dissipate. This is in stark contrast with the geometric model of adaptation (which is more relevant to cognitive skills), characterized in Proposition 2, for which the tension between short-term performance and long-term learning may be more severe.

Figure 4 Optimal Effort and Practice Intensity under the Multiplicative Model with Low Final Effort Requirement



Note. $h(w,b) = w^{\theta} b^{1-\theta}$, $\alpha = 0.7$, $\gamma = 0.8$, $\theta = 0.6$, $\epsilon = 0.9$, $\underline{w}_T = 0$, $b_1 = 1$, $\lambda = 0.9$, $\mu = 1.1$, $T = 12$, $\beta = 0.69$ (left) and $\beta = 0.9$ (right).

demonstrating the power of habit for skill retention, consistent with the empirical evidence of the benefits of lifelong exercise (Gries et al. 2018).

4.2.1. Low Final Effort Requirement. We first consider the case of a low final effort requirement, i.e., when the coupling constraint in $(P'_\times)$ is loose. The next proposition characterizes the optimal profile of the practice process and Figure 4 illustrates the optimal effort $\{z_t^*\}$ (solid lines) and practice intensity $\{w_t^*\}$ (dashed lines) both when $\alpha \geq \beta$ and when $\alpha < \beta$. Condition $(C_\times)$ below is obtained is obtained by first solving $(P'_\times)$ without the coupling constraint and then plugging that optimal solution into the coupling constraint.

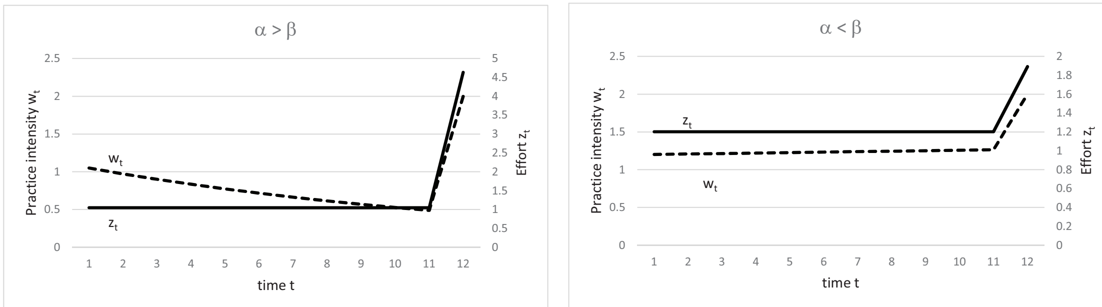
PROPOSITION 4. *For Problem $(P_\times)$, when the ultimate effort requirement is low, i.e., when the following condition holds:*

$$\zeta \cdot (h(\zeta, 1))^{T-1} \geq \frac{\underline{w}_T}{\epsilon^{T-1} b_1}, \qquad (C_\times)$$

in which ζ is the unique positive root of $\beta\lambda\zeta^{\lambda} - (\mu - \lambda)\zeta^{\lambda+\mu} - \alpha\mu\zeta^{\mu}$, the optimal effort $\{z_t^*\}$ is constant and equal to ζ and the optimal practice intensity $\{w_t^*\}$ is monotone.

In contrast to the additive model, which prescribes monotone effort (Proposition 1) for skill acquisition, the multiplicative model prescribes *constant* effort for skill retention. Moreover, because the optimal effort level (ζ) is independent of T, the time horizon does not matter for skill retention, provided that the practice process does not involve a high effort at its end. However, given that the base level changes over time, practice intensity $\{w_t\}$ may be decreasing or increasing, and the direction of change can be shown to depend on the relative rates of decay of fitness and fatigue. Specifically, $\{w_t\}$ is decreasing if $\alpha \gg \beta$ and increasing if $\alpha \ll \beta$.

Figure 5 **Optimal Effort and Practice Intensity under the Multiplicative Model with Geometric Adaptation and High Final Effort Requirement**



Note. $\alpha = 0.7$, $\gamma = 0.8$, $\theta = 0.6$, $\epsilon = 0.9$, $\underline{w}_T = 2$, $b_1 = 1$, $\lambda = 0.9$, $\mu = 1.1$, $T = 12$, $\beta = 0.69$ (left) and $\beta = 0.9$ (right).

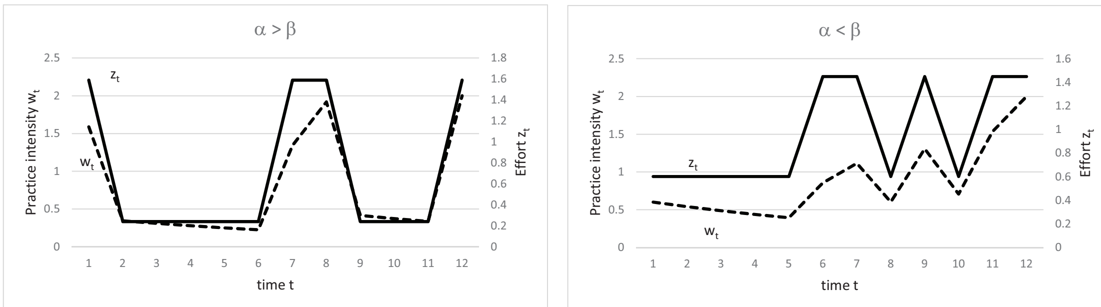
4.2.2. High Final Effort Requirement. We next consider the case of a high final effort requirement, i.e., when the coupling constraint in $(P'_\times)$ is tight at the optimum. We consider separately the cases of geometric and maximum adaptation. Overall, we find that the optimal practice process operates in cycles of either constant effort under geometric adaptation or pulsed effort under maximum adaptation.

Geometric Adaptation (Cognitive Tasks). We first consider the case of geometric adaptation, which, as discussed in §3.2.1, tends to fit better cognitive and some discrete motor tasks. The next proposition characterizes the optimal practice process and Figure 5 illustrates the optimal effort $\{z_t^*\}$ (solid lines) and practice intensity $\{w_t^*\}$ (dashed lines) both when $\alpha \geq \beta$ and when $\alpha < \beta$.

PROPOSITION 5. *For Problem $(P_\times)$ with $h(w,b) = w^\theta b^{1-\theta}$, when the ultimate effort requirement is high, i.e., when Condition $(C_\times)$ does not hold, the optimal effort $\{z_t^*\}_{t=1,\ldots,T-1}$ is constant and the optimal practice intensity $\{w_t^*\}_{t=1,\ldots,T-1}$ is monotone.*

We find that the optimal effort is *constant* up to period $T-1$. Hence, the higher final effort requirement distorts the effort profile only by making the last effort differ from the preceding efforts. (However, the value of the constant effort $\{z_t^*\}_{t=1,\ldots,T-1}$ is now dependent on the time horizon T.) Hence in settings where the multiplicative model with geometric adaptation is the most relevant, such as the retention of cognitive and some discrete motor skills, constant effort is optimal up to (but with the exclusion of) the final performance assessment. The resulting practice intensity $\{w_t^*\}_{t=1,\ldots,T-1}$ is monotone. Under a technical condition (see Corollary C-1 in Appendix EC.3 for details), we show that the practice profile $\{w_t^*\}$ depends on the decay rates of fitness and fatigue in a similar fashion to the case where efforts are unconstrained (Proposition 4). In sum, the tightening of the coupling constraint in $(P'_\times)$ raises the effort exerted throughout the practice process, and in particular the last-period's effort; however, the overall pattern of the practice process, with the exception of the last period, remains unchanged.

Figure 6 **Optimal Effort and Practice Intensity under the Multiplicative Model with Maximum Adaptation and High Final Effort Requirement**



Note. $\alpha = 0.7$, $\gamma = 0.8$, $\epsilon = 0.9$, $\underline{w}_T = 2$, $b_1 = 1$, $\lambda = 0.9$, $\mu = 1.1$, $T = 12$, $\beta = 0.69$ (left) and $\beta = 0.9$ (right).

Maximum Adaptation (Continuous Motor Tasks). We finally consider the case of maximum adaptation, which, as discussed in §3.2.1, tends to fit better tasks that have a low degree of forgetting, such as continuous motor tasks. The next proposition characterizes the optimal practice process and Figure 6 illustrates a typical solution in terms of effort $\{z_t^*\}$ (solid lines) and practice intensity $\{w_t^*\}$ (dashed lines) when both $\alpha \geq \beta$ and $\alpha < \beta$. Similar to the additive model with maximum adaptation, the optimal solution is characterized in terms of a partition $\mathcal{S}^* \subseteq \{1, \ldots, T\}$, including period $\{T\}$, such that $z_t^* > 1$ for all $t \in \mathcal{S}^* \setminus \{T\}$ and $z_t^* \leq 1$ for all $t \notin \mathcal{S}^*$.

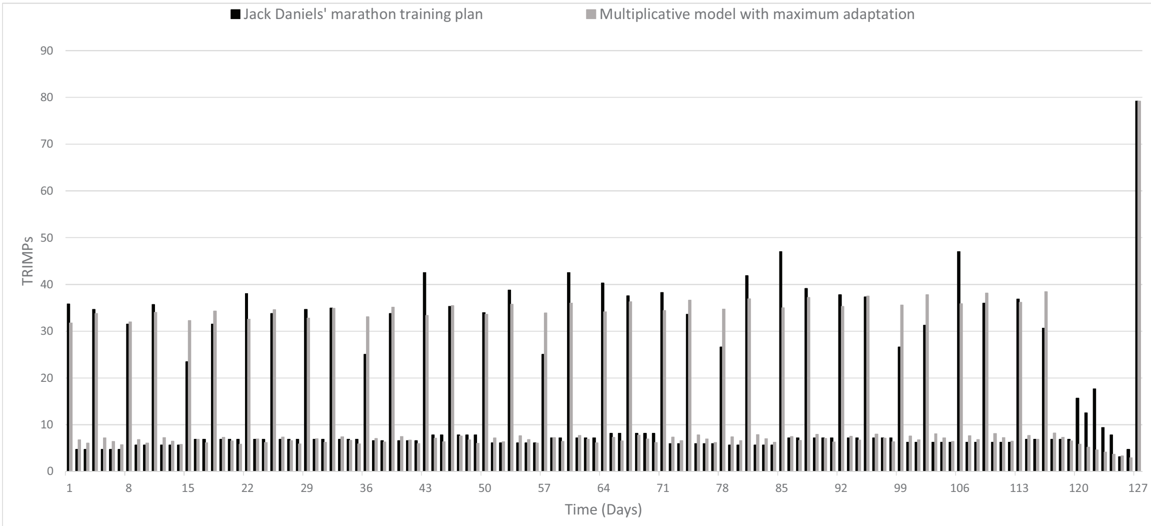
PROPOSITION 6. *For Problem ($P_\times$) with $h(w,b) = \max\{w, b\}$, when the ultimate effort requirement is high, i.e., when Condition ($C_\times$) does not hold, there exists an integer κ such that for any partition $\mathcal{S}^*$, $\{T\} \subseteq \mathcal{S}^* \subseteq \{1, \ldots, T\}$ with $|\mathcal{S}^*| = \kappa$,*

- *The optimal effort $\{z_t^*\}_{t \in \mathcal{S}^* \setminus \{T\}}$ is constant and equal to some $\overline{z} > 1$; across the high-effort sessions, the optimal practice intensity is monotone, i.e., for all t such that $\{t, t+1\} \subseteq \mathcal{S}^* \setminus \{T\}$, either $w_t \geq w_{t+1}$ or $w_t \leq w_{t+1}$;*
- *The optimal effort $\{z_t^*\}_{t \notin \mathcal{S}^*}$ is constant and equal to some $\underline{z} \leq 1$; within a low-effort session interval, the optimal practice intensity is decreasing, i.e., $w_t \geq w_{t+1}$ if $t+1 \notin \mathcal{S}^*$.*

Under maximum adaptation, $\{z_t^*\}$ may thus take on at most three values: $\overline{z}$, common across all high-effort sessions (up to $T-1$), $\underline{z}$, common across all low-effort sessions (up to $T-1$), and z_T^* in the last period. Hence, in contrast to the case with a low final effort requirement (Proposition 4) or the case with geometric adaptation (Proposition 5), effort is *pulsed.* To retain continuous motor skills (e.g., swimming, riding a bicycle) with the prospect of a challenging assessment, one thus needs to periodically alternate high-effort practices with low-effort practices, similar to the periodization principle in endurance sport training (Morton 1991, Smith 2003).

Note also the difference in effort patterns between the additive and the multiplicative models under maximum adaptation: Under the additive model, which fits better the skill acquisition phase

Figure 7 **Daniels' 18-Week Marathon Training Program and Fitted Multiplicative Model with Maximum Adaptation**



of learning, it is optimal to front-load the high-effort sessions (Proposition 3) so as to have time to recover or reset the base level before the ultimate session. In contrast, under the multiplicative model, which fits better the skill retention stage, high-effort practices can be periodically introduced to maintain a high base level b_t. As a result of this pulsed pattern of effort, the optimal practice intensity $\{w_t^*\}$ also follows a pulsed pattern.

Moreover, the optimal set of high-intensity practice sessions is only defined in terms of its size κ, not in terms of its actual composition. Accordingly, there exist multiple optimal solutions. For instance, the optimal value in the right panel of Figure 6 is obtained with the set $\mathcal{S}^* = \{6,7,9,11,12\}$, but it could have also been obtained with the set $\mathcal{S} = \{1,2,3,4,12\}$, which would have front-loaded the high-effort sessions, or the set $\mathcal{S} = \{8,9,10,11,12\}$, which would have back-loaded them. Hence, the optimal practice process offers some flexibility, and the choice of when to schedule the high-effort sessions may be driven by other considerations, such as the desire to build up a strong base early or the desire to develop a regular practice habit.

To illustrate the pulsed pattern under the multiplicative model with maximum adaptation, consider Daniels' 18-week marathon program, with 2 quality runs (i.e., runs combining long distance with speedwork) per week and a 55-mile peak mileage (Daniels 2014, Table 14.3). We convert the intensity of each session in Training Impulses (TRIMPs), weighting the recommended miles of each session by their recommended pace; see Appendix EC.2 for details. Figure 7 shows (in black) the intensities (in TRIMPs) of the daily practice sessions, including the marathon on the final day.

In addition, Figure 7 shows (in gray) the optimal practice process under the multiplicative model with maximum adaptation with the following parameter values: $\lambda = 0.999995$, $\mu = 1.0$,

$\alpha = 0.999995$, $\beta = 1.0$, $\epsilon = 0.999993$, and $b_1 = 32.7902$, which leads to an optimal value of low effort $\underline{z} = 0.194068$ and that of high effort $\overline{z} = 1.0041$. (See Appendix EC.2 for details on how these values were identified.) We observe that the difference in the estimated decay rates between fitness and fatigue, although central in the original fitness-fatigue model, is relatively marginal here. Similarly, the effect of nonlinearities seems small. As a result, the pulsed pattern of the practice process appears to be mostly driven by the adaptation mechanism, i.e., by the fact that effort adapts rapidly to high-intensity practices and decays slowly between practice periods.

In summary, the optimal practice process for skill retention (multiplicative model) may operate in cycles of either constant effort under geometric adaptation (typically, for cognitive tasks) and pulsed effort under maximum adaptation (typically, for continuous motor tasks), up to the last period, consisting of a final performance assessment.

5. Conclusion

This paper introduces a model of the practice process for skill acquisition and retention with applications to endurance sports training, motor learning, and cognitive learning. Adopting an optimization perspective, we characterize the optimal practice strategies that maximize performance on a given date. Our approach has business implications for sports analytics, training in organizations, and business model innovations in education.

Although we built on a model that was developed in the context of endurance sports training, practice process optimization can in principle be applied to other contexts, provided that the effect of practice on performance is fully understood. The model has the following characteristics. Each practice session affects performance positively by contributing to a stock of fitness and negatively by contributing to a stock of fatigue. Between practice sessions both fitness and fatigue decay exponentially. The effect of a practice on fitness and fatigue is defined in terms of effort, relative to the mean intensity of past practices. Furthermore, the effect is nonlinear, with decreasing marginal returns on fitness and increasing marginal costs on fatigue. Finally, the effect of practice on performance can be additive or multiplicative.

We characterize the optimal practice process under four variants of the model, depending on whether the effect of practice on performance is additive (corresponding to the skill acquisition stage) or multiplicative (corresponding to the skill retention stage), and depending on the effort adaptation mechanism to past practices, which is function of either their geometric mean (as would typically be the case for tasks that are easily forgotten with lack of practice such as cognitive tasks) or their maximum (as would typically be the case for tasks that are not easily forgotten with lack of practice such as continuous motor tasks). For each model variant, we consider two scenarios, namely when the ultimate effort requirement (when performance is assessed) is low or high.

Table 1 Optimal Effort Profiles

		Performance Model	
Final Effort Req.	Adaptation Mechanism	Additive (Skill Acquisition)	Multiplicative (Skill Retention)
Low		Monotone	Constant
High	Geometric (Cognitive Skills)	U-shaped (smooth)	Constant up to $T-1$
	Maximum (Cont. Motor Skills)	U-shaped (abrupt)	Pulsed

Table 1 summarizes the optimal effort profile for each model variant and scenario. For the additive model (skill acquisition), the optimal practice process involves *phases* of intensity increase and decrease, yielding U-shaped (or monotone) effort. Consistent with the principle of spacing out practices to enhance learning, practice intensity may first increase to build the base level and then decrease to either recover (if fatigue decays more quickly than fitness) or reset the base level (if fitness decays faster than fatigue); there may be one last increase in intensity near the end of the horizon, either limited to the final performance test or extending to the last set of practices preceding that performance assessment to rebuild fitness (if fitness decays faster than fatigue). The transitions in intensity are smooth under geometric adaptation (cognitive skills) but may be abrupt under maximum adaptation (continuous motor skills). In the latter case, it may be optimal to have a pronounced period of recovery before a demanding final performance test, similar to the tapering strategy in endurance sports.

For the multiplicative model (skill retention), the practice process consists of *cycles* of either constant effort under geometric adaptation (cognitive skills) or pulsed effort under maximum adaptation (continuous motor skills), consistent with the principle of alternating stress and rest.

Overall, our approach demonstrates the optimality of common high-performance practice strategies (and characterizes the conditions under which they are optimal) with a stylized model that involves a dual set of constructs, namely stocks (fitness and fatigue) and flow regulators (adaptation), formalizing similar ideas by Groves and Thompson (1970) and Bjork and Bjork (1992). Our findings suggest a key distinction between the practice process for skill acquisition, which proceeds in phases, and that for skill retention, which consists of cycles of either constant effort or pulsed effort. Our findings also suggest that the tension between short-term performance and long-term learning (Schmidt and Bjork 1992) is more severe for cognitive skills (geometric adaptation) than for continuous motor skills (maximum adaptation), since the practice process of the latter may include more pronounced periods of rest before a challenging assessment. Yet, multi-period breaks, which often arise in both research experiments on learning and practice (e.g., summer breaks), are typically suboptimal.

We outline several promising model extensions. First, the model could be extended to capture the multi-dimensional nature of learning, similar to Ladany (1975) and Zwols and Sierksma (2009), but with dynamic effects. In particular, interleaving different topics helps develop discrimination skills and is therefore more conducive to long-term retention (Brown et al. 2014), although there exist few guidelines about the optimal degree of interleaving. Second, it would be worthwhile to build a continuous-time version of the model, similar to Baucells and Zhao (2018), to accommodate non-impulse practices. Third, one might consider how to achieve longer-term performance objectives, such as health or long-term retention, with shorter-term objectives (e.g., races, exams). Finally, the profile of the practice process could be optimized subject to constraints on the total amount of practice (e.g., a 30-hour course) or on their duration (e.g., 1.5 hour) to shed light on the benefits and challenges of common course formats.

Given its parametric form, the proposed model can be tailored to fit individual characteristics and specific contexts. Our preliminary empirical analysis (Appendices EC.1 and EC.2) suggests that, in the context of long-distance running, adaptation matters more than the gap in decay rates between fitness and fatigue and other nonlinearities. Along these lines, future research should test the model's robustness with respect to its parameters and identify the typical values for those parameters in different contexts (motor vs. cognitive learning). A calibrated model could potentially be implemented in software to provide feedback to athletes, employees, or students about their performance over time and help them optimize their training or learning schedule. As the nature of work continues to evolve, requiring constant learning (Staats 2018), and as the development of technologies (such as wearables) provides greater access to tracking data on both the duration and intensity of practice and on performance, there is a growing opportunity for developing not only the science of training and learning, but also its engineering. We hope that this paper will inspire others to pursue this emerging field of research.

Acknowledgments

The author would like to thank Serguei Netessine (the DE), the review team, and Uday Karmarkar for helping strengthen the positioning and contribution of the paper. Thanks also to the seminar participants at the UCL School of Management, CORE (UC Louvain), and the Frankfurt School of Finance & Management.

References

Aflaki S, Popescu I (2013) Managing retention in service relationships. *Manage. Sci.* 60(2):415–433.

Anderson JR (2000) *Learning and Memory: An Integrated Approach* (New York, NY: Wiley), 2nd edition.

Atkinson RC (1972) Ingredients for a theory of instruction. *Am. Psychol.* 27(10):921.

Banister EW, Calvert TW, Savage MV, Bach TM (1975) A systems model of training for athletic performance. *Aust. J. Sports Med* 73(3):57–61.

Banister EW, Carter JB, Zarkadas PC (1999) Training theory and taper: Validation in triathlon athletes. *Eur. J. Appl. Physiol.* 79(2):182–191.

Baucells M, Sarin RK (2007) Satiation in discounted utility. *Oper. Res.* 55(1):170–181.

Baucells M, Sarin RK (2010) Predicting utility under satiation and habit formation. *Manage. Sci.* 56(2):286–301.

Baucells M, Zhao L (2018) It is time to get some rest. *Manage. Sci.* Articles in Advance.

Bernbach HA (1965) A forgetting model for paired-associate learning. *J. Math. Psychol.* 2(1):128–144.

Bjork RA, Bjork EL (1992) A new theory of disuse and an old theory of stimulus fluctuation. *From Learning Processes to Cognitive Processes: Essays in Honor of William K. Estes*, volume 2, chapter 2, 35–67 (Erlbaum, NJ: Hillsdale).

Bompa TO, Haff GG (2009) *Periodization Training in Sports* (Champaign, IL: Human Kinetics).

Bourne LE Jr, Archer EJ (1956) Time continuously on target as a function of distribution of practice. *J. Exp. Psychol.* 51(1):25–33.

Brown PC, Roediger HL, McDaniel MA (2014) *Make It Stick: The Science of Succesful Learning* (Cambridge, MA: Harvard University Press).

Busso T, Candau R, Lacour JR (1994) Fatigue and fitness modelled from the effects of training on performance. *Eur. J. Appl. Physiol.* 69(1):50–54.

Busso T, Denis C, Bonnefoy R, Geyssant A, Lacour JR (1997) Modeling of adaptations to physical training by using a recursive least squares algorithm. *J. Appl. Phys.* 82(5):1685–1693.

Busso T, Häkkinen K, Pakarinen A, Carasso C, Lacour JR, Komi PV, Kauhanen H (1990) A systems model of training responses and its relationship to hormonal responses in elite weight-lifters. *Eur. J. Appl. Physiol.* 61(1):48–54.

Calvert TW, Banister EW, Savage MV, M BT (1976) A systems model of the effects of training on physical performance. *IEEE T. Syst. Man Cyb.* 6(2):94–102.

Cepeda NJ, Pashler H, Vul E, Wixted JT, Rohrer D (2006) Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychol. Bull.* 132(3):354–380.

Chambliss DF (1989) The mundanity of excellence. an ethnographic report on stratification and Olympic swimmers. *Sociol. Theor.* 7(1):70–86.

Chan JCK, McDermott KB, Roediger HL III (2006) Retrieval-induced facilitation: initially nontested material can benefit from prior testing of related material. *J. Exp. Psychol. Gen.* 135(4):553.

Coggan A (2008) The science of the performance manager. URL `https://www.trainingpeaks.com/blog/the-science-of-the-performance-manager/`, last accessed on March 30, 2017.

Collignon H, Sultan N, Santander C (2011) The sports market. URL `https://www.atkearney.com/documents/10192/6f46b880-f8d1-4909-9960-cc605bb1ff34`, last accessed on November 15, 2018.

Crowder RG (1976) *Principles of Learning and Memory* (Hillsdale, NJ: Erlbaum).

Cuddy LJ, Jacoby LL (1982) When forgetting helps memory: An analysis of repetition effects. *J. Verb. Learn. Verb. B.* 21(4):451–467.

Dai H, Milkman KL, Hofmann DA, Staats BR (2015) The impact of time at work and time off from work on rule compliance: The case of hand hygiene in health care. *J. Applied Psychol.* 100(3):846.

Daniels J (2014) *Daniels' Running Formula* (Champaign, IL: Human Kinetics), 3rd edition.

Das Gupta A, Karmarkar US, Roels G (2015) The design of experiential services with acclimation and memory decay: Optimal sequence and duration. *Manage. Sci.* 62(5):1278–1296.

Davenport TH (2014) What businesses can learn from sports analytics. *MIT Sloan Manage. Rev.* 55(4):10–13.

Ebbinghaus H (1913) *Memory: A Contribution to Experimental Psychology* (New York, NY: Teachers College, Columbia University), (Original work published 1885).

Ericsson KA, Krampe RT, Tesch-Römer C (1993) The role of deliberate practice in the acquisition of expert performance. *Psychol. Rev.* 100(3):363.

Eysenck HJ (1965) A three-factor theory of reminiscence. *B. J. Psychol.* 56(2-3):163–182.

Fisher M, Gallino S, Netessine S (2018) Does online training work in retail?, URL `https://dx.doi.org/10.2139/ssrn.2670618`, working paper, the Wharton school.

Fitze-Clarke JR, Morton RH, Banister EW (1991) Optimizing athletic performance by influence curves. *J. Appl. Physiol.* 71(3):1151–1158.

Gerchak Y (1994) Operations research in sports. Pollock SM, Rothkopf MH, Barnett A, eds., *Operations Research and the Public Sector*, volume 6 of *Handbooks in Operations Research and Management Science*, chapter 14, 507–527 (Elsevier).

Gladwell M (2008) *Outliers: The Story of Success* (Hachette UK).

Green LV, Savin S, Savva N (2013) "Nursevendor problem": Personnel staffing in the presence of endogenous absenteeism. *Manage. Sci.* 59(10):2237–2256.

Gries KJ, Raue U, Perkins RK, Lavin KM, Overstreet BS, D'Acquisto LJ, Graham B, Finch WH, Kaminsky LA, Trappe TA, et al. (2018) Cardiovascular and skeletal muscle health with lifelong exercise. *J. Appl. Physiol.* 125(5):1636–1645.

Groves PM, Thompson RF (1970) Habituation: a dual-process theory. *Psychol. Rev.* 77(5):419.

Hardy G, Littlewood JE, Pólya G (1952) *Inequalities* (New York, NY: Cambridge University Press), 2nd edition.

Hax AC, Majluf NS (1982) Competitive cost dynamics: the experience curve. *Interfaces* 12(5):50–61.

Heathcote A, Brown S, Mewhort DJK (2000) The power law repealed: The case for an exponential law of practice. *Psychon. B. Rev.* 7(2):185–207.

Hellard P, Avalos M, Millet G, Lacoste L, Barale F, Chatard JC (2005) Modeling the residual effects and threshold saturation of training: A case study of Olympic swimmers. *J. Strength Cond. Res.* 19(1):67–75.

Henderson BD (1984) The application and misapplication of the experience curve. *Journal of Business Strategy* 4(3):3–9.

Hull CL (1943) *Principles of Behavior*. The Century Psychology Series (New York, NY: Appleton Century Crofts).

Ibanez MR, Clark JR, Huckman RS, Staats BR (2017) Discretionary task ordering: Queue management in radiological services. *Manage. Sci.* 64(9):4389–4407.

Jaber MY, Bonney M (1997) A comparative study of learning curves with forgetting. *Appl. Math. Model.* 21(8):523–531.

Karmarkar US, Karmarkar UR (2014) Customer experience and service design. Baglieri E, Karmarkar US, eds., *Managing Consumer Services. Factory or Theater?*, chapter 7, 109–130 (Springer).

KC DS, Terwiesch C (2009) Impact of workload on service time and patient safety: An econometric analysis of hospital operations. *Manage. Sci.* 55(9):1486–1498.

Kraepelin E (1902) Die arbeitscurve. *Philosophische Studien* 19:459–507.

Kuper S, Szymanski S (2014) *Soccernomics: Why England Loses, Why Spain, Germany, and Brazil Win, and Why the US, Japan, Australia and Even Iraq Are Destined to Become the Kings of the World's Most Popular Sport* (Nation Books).

Ladany SP (1975) Optimization of pentathlon training plans. *Manage. Sci.* 21(10):1144–1155.

Lewis M (2004) *Moneyball: The Art of Winning an Unfair Game* (WW Norton & Company).

Machol RE, Ladany SP, Morrison DG, eds. (1976) *Management Science in Sports* (Elsevier/North Holland, Inc.).

Morton RH (1991) The quantitative periodization of athletic training: A model study. *Res. Sports Med.* 3(1):19–28.

Morton RH, Fitze-Clarke JR, , Banister EW (1990) Modeling human performance in running. *J. Appl. Physiol.* 69(3):1171–1177.

Moxnes JF, Hausken K (2008) The dynamics of athletic performance, fitness and fatigue. *Math. Comp. Model. Dyn.* 14(6):515–533.

Mujika I, Padilla S (2003) Scientific bases for precompetition tapering strategies. *Med. Sci. Sport Exer.* 35(7):1182–1187.

Nasiry J, Popescu I (2011) Dynamic pricing with loss-averse consumers and peak-end anchoring. *Oper. Res.* 56(6):1361–1368.

Nerlove M, Arrow KJ (1962) Optimal advertising policy under dynamic conditions. *Econometrica* 29(114):129–142.

Newell A, Rosenbloom PS (1981) Mechanisms of skill acquisition and the law of practice. Anderson JR, ed., *Cognitive Skills and their Acquisition*, volume 1, 1–55 (Hillsdale, NJ: Erlbaum).

Newell A, Simon HA (1972) *Human Problem Solving* (Prentice-Hall Englewood Cliffs, NJ).

Popescu I, Wu Y (2007) Dynamic pricing strategies with reference effects. *Oper. Res.* 55(3):413–429.

Powell A, Savin S, Savva N (2012) Physician workload and hospital reimbursement: Overworked physicians generate less revenue per patient. *M&SOM Manuf. Serv. Op.* 14(4):512–528.

Roediger HL III, Karpicke JD (2006) Test-enhanced learning: Taking memory tests improves long-term retention. *Psychol. Sci.* 17(3):249–255.

Schmidt RA, Bjork RA (1992) New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychol. Sci.* 3(4):207–218.

Schmidt RA, Lee TD (2011) *Motor Control and Learning: A Behavioral Perspective* (Champaign, IL: Human Kinetics), 5th edition.

Simon DA, Bjork RA (2001) Metacognition in motor learning. *J. Exp. Psychol. Learn.* 27(4):907–912.

Simon H (1982) ADPULS: An advertising model with wearout and pulsation. *J. Marketing Res.* 19(3):352–363.

Smith DJ (2003) A framework for understanding the training process leading to elite performance. *Sports Med.* 33(15):1103–1126.

Sonnentag S, Kleine BM (2000) Deliberate practice at work: A study with insurance agents. *J. Occup. Organ. Psych.* 73(1):87–102.

Staats BR (2018) *Never Stop Learning: Stay Relevant, Reinvent Yourself, and Thrive* (Harvard Business Press).

Staats BR, Gino F (2012) Specialization and variety in repetitive tasks: Evidence from a Japanese bank. *Manage. Sci.* 58(6):1141–1159.

Strauss V (2013) Global education market reaches $4.4 trillion – and is growing. *The Wall Street Journal* URL `https://www.washingtonpost.com/news/answer-sheet/wp/2013/02/09/global-education-market-reaches-4-4-trillion-and-is-growing/`, february 9.

Stulberg B, Magness S (2017) *Peak Performance* (New York, NY: Rodale).

Taha T, Thomas SG (2003) Systems modelling of the relationship between training and performance. *Sports Med.* 33(14):1061–1073.

Tan TF, Netessine S (2014) When does the devil make work? an empirical study of the impact of workload on worker productivity. *Manage. Sci.* 60(6):1574–1593.

Taylor FW (1911) *The Principles of Scientific Management* (Harper).

Terry WS (2009) *Learning and Memory* (Boston, MA: Pearson Education), 4 edition.

Terwiesch C, Ulrich KT (2014) Will video kill the classroom star? The threat and opportunity of massively open online courses for full-time MBA students, Mack Institute, the Wharton school.

The Economist (2017) Lifelong learning is becoming an economic imperative. January 12.

Thorndike EL (1912) The curve of work. *Psychol. Rev.* 19(3):165.

Wallace LK, Slattery KM, Coutts AJ (2014) A comparison of methods for quantifying training loads: relationships between modelled and actual training responses. *Eur. J. Appl. Physiol.* 114(1):11–20.

Wathieu L (1997) Habits and the anomalies in intertemporal choice. *Manage. Sci.* 43(11):1552–1563.

Williams O (1926) A study of the phenomenon of reminiscence. *J. Exp. Psychol.* 9(5):368.

Wixted JT, Carpenter SK (2007) The Wickelgren power law and the Ebbinghaus savings function. *Psychol. Sci.* 18(2):133–134.

Wood RE, Hayter S, Rowbottom D, Stewart I (2005) Applying a mathematical model to training adaptation in a distance runner. *Eur. J. Appl. Physiol.* 94(3):310–316.

Wright TP (1936) Factors affecting the cost of airplanes. *J. Aeronaut. Sci.* 3(4):122–128.

Xu Y, Tan T, Netessine S (2017) When is the root of all evil not money? The impact of load on operational risk at a commercial bank, URL `https://ssrn.com/abstract=3075452`, Wharton working paper.

Zwols Y, Sierksma G (2009) Training optimization for the decathlon. *Oper. Res.* 57(4):812–822.