MODELS OF FATIGUE AND REST IN LEARNING. PART II: Motor and Cognitive Skills

D. I. Grebenkov*, A. A. Kozlova**, D. V. Lemtyuzhnikova***, and D. A. Novikov****

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

*✉ grebenkov-d-i@mail.ru, **✉ sankamoro@mail.ru, ***✉ darabbt@gmail.com, ****✉ novikov@ipu.ru

Abstract. This paper considers the possibilities and examples of using a general learning model with fatigue and rest effects to describe experimental data. A classification of iterative learning models is provided, and the existing datasets on learning from various fields are overviewed. A general algorithm for selecting an appropriate iterative learning model based on available experimental data is proposed. Examples of processing experimental and modeling data are presented for motor and cognitive skills, visual-motor adaptation, and tasks with long breaks. The following hypothesis is formulated and tested: learning models describe the data with deviations representing independent and identically distributed realizations of Gaussian random variables with zero mean. According to the values of statistical criteria, there are no grounds to reject this hypothesis. Based on the modeling results, recommendations on the optimization and management of the learning process are given.

Keywords: experience, iterative learning, learning curve, mathematical modeling, fatigue, rest, skill acquisition, forgetting, experiment.

INTRODUCTION

This paper is part II of the study devoted to extending the general learning model [1, 2] by including the processes of fatigue and rest. In part I of the study, we reviewed the main approaches to describing the learning process as well as supplemented classical mathematical models of experience acquisition [3]. Let us recall the key concepts and definitions discussed in detail in [3].

Learning, the process and result of acquiring individual experience [1], underlies the adaptation of living and nonliving systems to changing conditions. In the context of mathematical modeling, learning is understood as a process during which a (biological, technical, or abstract-logical) system optimizes its actions to achieve a given goal. Of particular interest is *iterative learning* (IL), a type of learning based on the system's repetition of actions (trials and errors) to achieve a fixed goal under constant external conditions [4]. This process is the foundation for developing skills in humans, conditioned reflexes in animals, and adaptation algorithms in robotics and artificial intelligence.

Mathematical models of IL describe a sequence of learning levels (the so-called *learning curves*, LCs) through systems of equations, graphs, or algorithms to reveal universal regularities. For example, an LC is the probability of mastering an activity component depending on time or the number of repetitions (iterations).

The classical exponential LC has the form

$$q(t) = q_{\max} - (q_{\max} - q_{\min}) e^{-\gamma t},$$

where $q(t)$ is the current learning level; $q_{\max}$ is the highest learning level observed; $q_{\min}$ is the initial learning level; $t \ge 0$ is time; finally, $\gamma > 0$ is the learning rate.

Based on experimental data, it can be concluded that these curves have a slowly asymptotic nature: the rate of performance improvement decreases with time, and the curve tends to some limit [4, 5]. Such LCs are often approximated by exponential functions [2, 4].

Part I of the study considered approaches to learning modeling and presented extended versions of the models from [1, 4]. In this paper, we apply the extended models to the processing of real data and provide an algorithm for processing such data (see Section 2).

1. A REVIEW OF EXPERIMENTAL DATA AND CLASSIFICATION OF MODELS

To model the learning process, one needs data reflecting experience (skill) acquisition dynamics with the influence of fatigue, the role of rest, and the individual characteristics of learners. This paper considers 11 datasets covering various fields, ranging from educational platforms to sports training. These data allow identifying the key processes (mastering, forgetting, fatigue, and rest) and parameters (time, history, and fatigue rate) necessary for building learning models.

Each dataset is unique in its structure and context, but all datasets have a common objective: to collect information on how people learn, fatigue, rest, and forget, and to make this information available to a wide range of researchers. Some datasets, such as Duolingo Spaced Repetition Data [6] and ASSISTments [7], focus on cognitive aspects, while others, such as Motor Learning Experiments [8] and Injury Prediction in Competitive Runners [9], explore motor skills and physical fatigue.

The information about the processes and parameters of each dataset is systematized in the table below (see the Appendix), including common regularities that may be useful for building learning models. Based on the analysis of the 11 datasets, we identify the key processes and parameters that should be considered when modeling learning.

These processes include:

• *Mastering* is the central process present in all datasets. This process manifests itself in improved results over time: in fewer errors (ASSISTments, Junyi Academy Online Learning Activity [10]), increased accuracy (Motor Learning Experiments), or reduced task completion time [11].

• *Forgetting* is observed in most datasets, especially in educational (Duolingo Spaced Repetition Data, ASSISTments) and motor (Motor Learning Experiments) datasets. This process manifests itself in decreased accuracy or increased errors after periods of inactivity.

• *Fatigue* is an important factor affecting performance in long sessions. This process manifests itself through an increase in errors (ASSISTments), an increase in task completion time (Codeforces Users Submissions [12]), or subjective assessments of fatigue (Injury Prediction in Competitive Runners).

• *Rest* plays a key role in skill recovery. In datasets, it is represented as periods of inactivity (Duolingo Spaced Repetition Data, Junyi Academy Online Learning Activity) or days without training (Injury Prediction in Competitive Runners).

• *The choice of skill acquired* is less explicitly represented; however, it can be traced in datasets where tasks vary in difficulty (Codeforces Users Submissions) or require adaptation to changing conditions (GradualTwoRate [13]).

The parameters and factors considered when calculating the current learning level are:

• *Time* is a universal parameter present in all datasets. It allows one to track the dynamics of learning, forgetting, and fatigue.

• *The history* of actions (correct and incorrect responses, movement trajectories) helps to understand how past experiences influence current results.

• *The number of skills acquired*.

• *Fatigue rate* is measured through objective metrics (errors, completion time) or subjective assessments (participant surveys).

These conclusions bring us to the next step in the development of learning models: using data from different sources to create unified approaches to modeling learning processes in various fields of human activity.

2. PROCESSING PRINCIPLES FOR EXPERIMENTAL DATA

Various characteristics are used as criteria for the learning level to describe the learning process of real agents and systems. According to [4], they are divided into the following categories:

– temporal (the time to perform/complete an action, operation, or task; reaction time; error correction time);

– speed-related (the quantities inverse to time, such as reaction speed, movement speed, or labor productivity, e.g., the number or share of correctly performed actions, the volume of production (output) per unit of time);

– accuracy-related (the error measured in physical quantities, the number of errors, or their probability);

– informational (the amount of information processed in a given time; the amount of material perceived or memorized).

In this paper, we analyze experimental data with accuracy-related and temporal characteristics used as criteria for the learning level. As shown in [5], regardless of the characteristic selected (temporal, speed- and accuracy-related, or informational), the LC has a slowly asymptotic nature. For speed-related characteristics, this type of curve will be achieved under a transformation replacing the inverse dependence on time with a direct one.

Another classification is also possible, based on the units of measurement of the learning level:

- absolute:
  - – units (the number of correctly or incorrectly performed actions),
  - – time (in s),
  - – speed (in 1/s),
  - – angle (in rad),
  - – length (in m),
  - – mass (in kg),
  - – information volume (in bits), etc.;
- relative:
  - – the share of correctly or incorrectly performed actions,
  - – the probability of a correct answer,
  - – the normalized values of absolute units.

Note that the vast majority of mathematical learning models operate relative (usually *probabilistic* [1]) characteristics of the learning level; see the reviews in [2, 4]. It is possible to pass from absolute values to relative ones by normalizing the range of learning levels ensuring the required nature (increase or decrease) of the LC:

$$\hat{q}(t) = \frac{q(t) - q_{\min}}{q_{\max} - q_{\min}} \text{ or } \hat{q}(t) = \frac{q_{\max} - q(t)}{q_{\max} - q_{\min}}, \quad (1)$$

where $\hat{q}(t)$ is the normalized estimate of the learning level.

The transformation (1) is linear; therefore, in contrast to threshold transformations [5], it preserves the qualitative properties of LCs. Also, if the initial LC contains some additive random noise, i.e.,

$$q_\theta(t) = q(t) + \theta_t, \quad (2)$$

where $q_\theta(t)$ is the "noisy" learning level and $\{\theta_t\}$ are independent and identically distributed (i.i.d.) realizations of a random variable with zero mean, the transformation (1) will not change the mathematical expectation of the noise, and the linear (normalizing) transformation will preserve the properties of both the LC and the additive "noise."

The experience acquisition models considered in [1, 2] are "*binary*": the learner's state with respect to an uncertainty factor (the component of the activity technology) takes two possible values, "1" (acquired) or "0" (not acquired, forgotten). Therefore, they use mathematical expectations (for the number of acquired technology components, uncertainty factor values, etc.) as the learning level (the experience criterion), leading to "continuous" LCs of type 0. Such LCs describe almost any practically interpretable situations and/or experimental data.

The LC of type 0 has many useful properties. For simplicity, let $q_{\max} = 1$ and $q_{\min} = 0$:

$$q(t) = 1 - e^{-\gamma t}.$$

Based on available experimental results, an appropriate IL model is selected by the following **algorithm**:

1. Verify that the learning process in the experiment is iterative, i.e., the learning system, the goal of learning, and the learning conditions are invariable.
2. Analyze the experimental data, identify the criterion for the learning level, and find its place in the above classification system (see the beginning of Section 2).
3. If necessary, perform the normalization (1).
4. Analyze, from a practical point of view, which of the IL models considered in [1–4] best matches the experiment.
5. Analyze which of the IL models best approximates the experimental data (using the least squares method or another method, particularly under the representation (2) with zero mean of the additive noise obeying the Gaussian or other symmetric distribution).
6. Analyze the practical interpretations of the identified parameters; in the case of several learners and/or series of experiments, compare them.
7. Compare the results of the selected model with the corresponding "optimal" learning conditions, e.g., with the optimal duration of a break in terms of maximizing the terminal learning level (see Part I of the study [3]). Formulate recommendations for optimizing the learning process.

Let us illustrate the application of this algorithm.

3. EXAMPLES OF EXPERIMENTAL DATA PROCESSING

Various statistical criteria are used to check whether the model deviations from the experimental data obey the Gaussian (normal) distribution. In this paper, we select the Shapiro–Wilk, Epps–Pulley [14], and Kolmogorov–Smirnov [15] tests for normality.

The *scipy.stats* module of *Python* was used to calculate the values of the Shapiro–Wilk and Kolmogorov–Smirnov test statistics. An extended version of the Shapiro–Wilk test for samples of up to 5000 elements was taken. The values of the Epps–Pulley test statistics were calculated by the software implementation [16] based on the test proposed in [17]. The tabulated values were borrowed from [18] and [19] (the Shapiro–Wilk test), [20] (the Kolmogorov–Smirnov test), and

[14] and [21] (the Epps–Pulley test). The *curve_fit* method from the *scipy.optimize* module was applied to estimate the model parameters, with the Levenberg–Marquardt minimization algorithm.

3.1. The Ball-Tossing Experiment

Consider the data of the experiment described in [22]: throwing balls into a box. The ultimate objective of the experiment was to study the learning capability by trial and error when learning is spread over a significant period of time.

A total of 200 ball-throwing attempts were made in one training day. The balls that fell into the box were scored. Daily training sessions were held at 9:00 a.m. Sunday was a day of rest. The experiment lasted 100 days. The results are shown in Fig. 1.

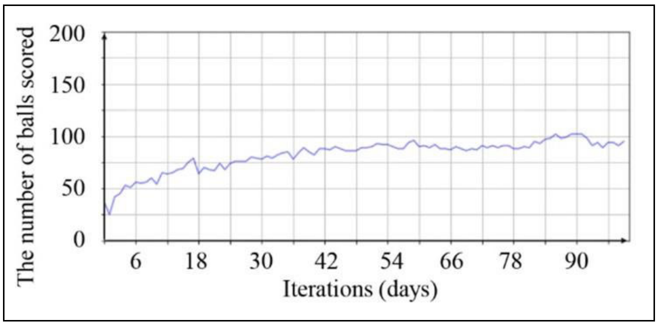

Fig. 1. The experimental data [22].

During the first 20–30 iterations, a marked improvement in the results can be observed, indicating rapid skill formation. Further on, the curve stabilizes, which demonstrates saturation of the learning process. This regularity is common for individual experience acquisition processes.

In the case of the single skill under consideration, the value $q(t)$ of the individual experience criterion (the so-called learning level) of an agent (trainee) is the probability that his/her experience will be formed in period t and not forgotten [3]. Let the criterion for the learning level $q(t)$ be the average frequency of scoring a ball on day t, which corresponds to the application of the transformation (1). According to the above classification, this is a relative accuracy characteristic (a quantity "inverse" to the probability of error, i.e., missing the box when throwing the ball). Let us construct an LC based on the experimental data (Fig. 2).

This curve has an asymptote, slightly deviating from the latter over time. We take two simple models from part I of the study [3] to reflect such dynamics.

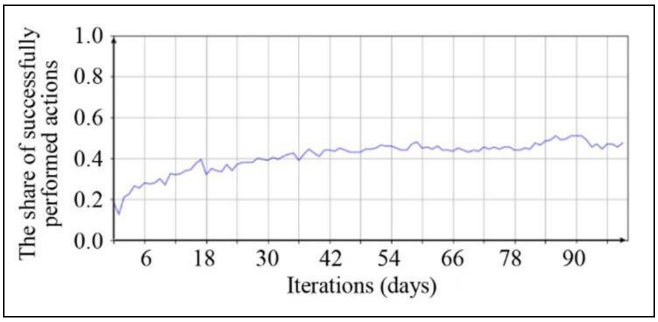

Fig. 2. The empirical LC based on the experimental data [22].

Within Model 3, mastering and forgetting occur independently of time [3]:

$$q(t+1) = q(t) + (1 - q(t))w - q(t)(1-u), \tag{3}$$

$$q(t) = \frac{w}{w+u} - \left(\frac{w}{w+u} - q(0)\right)e^{-(w+u)t}, \tag{4}$$

where w is the probability of acquiring a skill on iteration t; u is the probability of forgetting a learned skill on iteration t; finally, $q(0)$ is the initial skill level.

Under the condition $q(0) < \frac{w}{w+u}$, the curve (4) is nondecreasing and asymptotically tends to $\frac{w}{w+u}$. We approximate the data in Fig. 2 by the curve (4), selecting the values of w, u, and $q(0)$ so as to minimize the sum of squared deviations. Numerical estimation gives the following values of the model parameters: $q(0) \approx 0.187$, $w \approx 0.022$, and $u \approx 0.025$. The theoretical asymptote is $y = 0.468$. With the above parameters substituted into the model, the graph of the learning level is shown in Fig. 3.

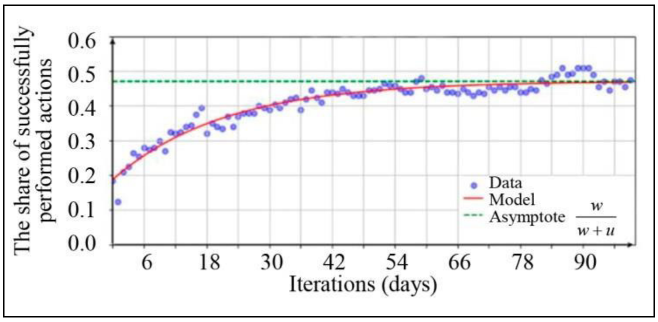

Fig. 3. The approximation of the experimental data [22] using model (4).

The LC in Fig. 3 has a slowly asymptotic nature and tends to the asymptote $y = 0.468$. With a known asymptote, one can predict the achievable skill level and optimize the duration of the learning process, i.e., terminate it when reaching a value sufficiently close to this level.

Suppose that Model (4) matches the experimental data with an additive random error due to data “noise” or some independent random factors in the course of the experiment:

$$q_\theta(t) = q(t) + \theta_t,$$

where θ_t are i.i.d. random variables obeying the Gaussian distribution with zero mean.

Figure 4 provides the histogram of the deviations between the predictions of model (4) and the experimental data $q_\theta(t)$.

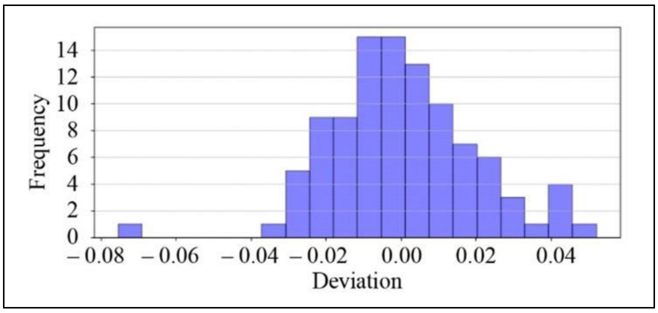

Fig. 4. The histogram of deviations of model (4) from the experimental data [22].

According to Fig. 4, the largest number of the deviations $q_\theta(t)$ is near zero, and the distribution visually resembles a Gaussian (normal) one. We hypothesize that the distribution of the deviations is Gaussian:

$$\theta \sim N(\mu, \sigma^2). \quad (5)$$

Let us verify this hypothesis using the Shapiro–Wilk test. The test statistic value is 0.972, and the p-value is 0.029 (with a significance level of 0.05). This means that the distribution is not Gaussian. At the same time, the Epps–Pulley and Kolmogorov–Smirnov tests have values of 0.298 and 0.059, respectively, and when compared with the critical values, they give no grounds to reject the normality hypothesis with the same significance level.

Now we repeat similar operations for Model 4 [3]:

$$q(t+1) = q(t) + w(t)\big(1 - q(t)\big), \quad (6)$$

$$q(t) = 1 - \big(1 - q(0)\big)e^{-W(t)}, \quad (7)$$

with the following estimate for the probability of mastering:

$$w(t) = w_0 e^{-\alpha t}, \quad (8)$$

$$W(t) = \frac{w_0}{\alpha}\left(1 - e^{-\alpha t}\right). \quad (9)$$

Let us approximate the curve in Fig. 2 by selecting the values of the parameters $q(0)$, w_0, and α for model (7)–(9): $q(0) \approx 0.1828$, $w_0 \approx 0.0179$, and $\alpha \approx 0.0407$. The graph of the learning level is presented in Fig. 5.

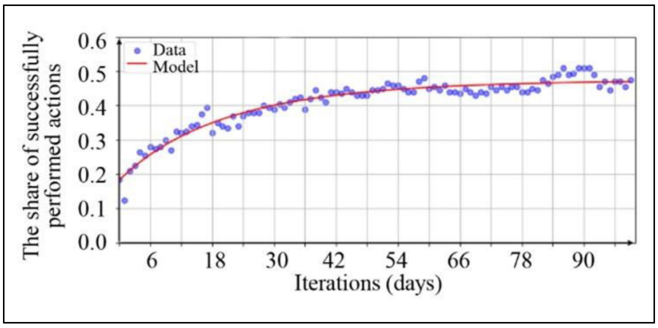

Fig. 5. The approximation of the experimental data [22] using model (7)–(9).

The experimental data demonstrate a rapid improvement in the results until approximately the 30th day, after which growth decelerates and gradually plateaus (reaches a plateau) at a learning level of approximately 0.5. Model (7)–(9) accurately describes these dynamics: the initial growth is ensured by a high probability of learning, and the subsequent slowdown is ensured by its exponential decrease.

Figure 6 shows the probability of experience acquisition $w(t)$ depending on time.

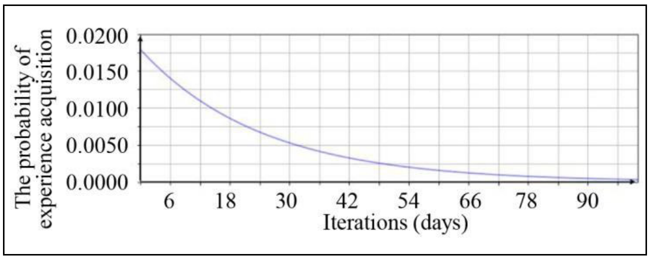

Fig. 6. The probability of experience acquisition depending on time in model (7)–(9) for the experimental data [22].

At the beginning of learning, $w(0) \approx 0.018$, but the probability of experience acquisition falls below 0.01 by the 20th day and below 0.005 by the 50th day; and it approaches zero by the end of the experiment. This exponential decrease describes the phenomenon of reduced learning level growth over time.

To assess the approximation quality of model (4), we analyze the distribution of the deviations between its predictions and the factual values of the learning level. The corresponding histogram can be found in Fig. 7.

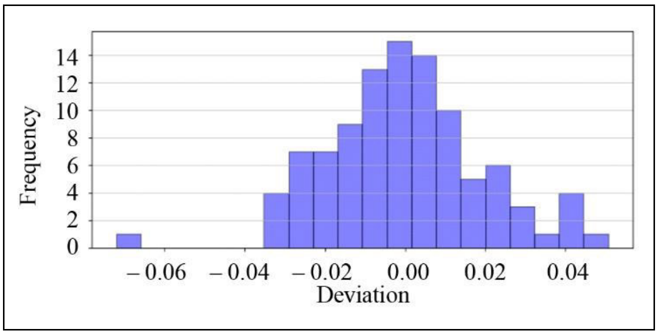

Fig. 7. The histogram of deviations of model (7)–(9) from the experimental data [22].

Most of the deviations range from –0.02 to 0.02, the distribution is visually symmetrical and centered around zero. This gives reason to assume that the model has no significant bias and matches the experimental data.

Let us verify hypothesis (5) using the Shapiro–Wilk test. The test statistic value is 0.977: for the distribution of the deviations with the parameters $\mu \approx 0$ and $\sigma \approx 0.02$, the normality hypothesis with a significance level of 0.05 cannot be rejected. According to the Epps–Pulley and Kolmogorov–Smirnov tests, this hypothesis is also not rejected. Therefore, we consider the data to be normally distributed. Hence, the deviations of the model predictions from the factual data can be supposed random without systematic errors.

Thus, model (7)–(9) matches the experimental data [22]. In addition, the deviations obey the Gaussian distribution with zero mean. As a result, two of the three tests confirm the normality of the deviations for model (4); for model (7)–(9), normality is confirmed by all three tests.

The above experiment on acquiring a motor skill (ball-tossing) reflects the iterative process of experience acquisition. The empirical curve demonstrates a smooth growth with an asymptote, typical behavior for slowly asymptotic LCs. According to the approximation results of the experimental data, the models from part I of the study [3] reproduce both the initial phase of active learning and the saturation phase. The deviations from the empirical data have the Gaussian distribution with an almost zero mean. The dynamics of the learning level confirm the need to control the duration of the learning process. For instance, at the phase of decelerated efficiency growth, it may be advisable to change the learning strategy or even terminate this process when reaching a learning level sufficiently close to the horizontal asymptote of the LC.

3.2. The Mental Multiplication Experiment

In this subsection, we consider the following experiment from the paper [11]: a participant solved a module of 63 exercises on a daily basis, multiplying two two-digit numbers in the brain. The time taken to complete the module was recorded in minutes. Figure 8 shows the time taken to complete the task depending on the number of modules solved (days).

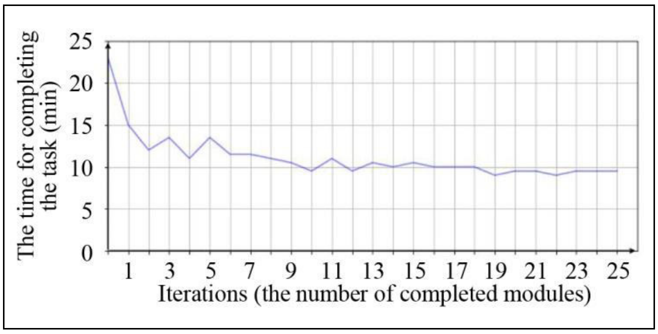

Fig. 8. The experimental data [11].

At the beginning of the experiment, the time fluctuates between 15 and 20 mins, then gradually decreases and, after the 15th iteration, stabilizes at 9–10 mins. A plateau is visible: the subsequent training leads to an insignificant acceleration. In other words, by the 15th–17th iteration, skill acquisition almost reaches its maximum, and continuing training in the previous conditions becomes inefficient.

To describe the processes of experience acquisition based on these experimental data, the models from part I of the study [3] have to be transformed into an increasing slowly asymptotic nature, with the values of the learning level belonging to the range [0, 1]. We convert the temporal absolute values of the learning level criterion into relative ones characterizing the proximity to the minimum possible time for completing the task:

$$\hat{q}(t) = \frac{\tau_{\max} - \tau(t)}{\tau_{\max} - \tau_{\min}}, \tag{10}$$

where $\tau_{\min}$ and $\tau_{\max}$ are the minimum and maximum task completion times, respectively; $\tau(t)$ is the task completion time on iteration *t*. (Also, see the more general transformation (1).)

For $\tau_{\min}$ = 5 mins and $\tau_{\max}$ = 25 mins, let us construct an LC based on the obtained data (Fig. 9).

By the 3rd or 4th iteration, the degree of proximity exceeds 0.6, and starting from the 10th iteration, it

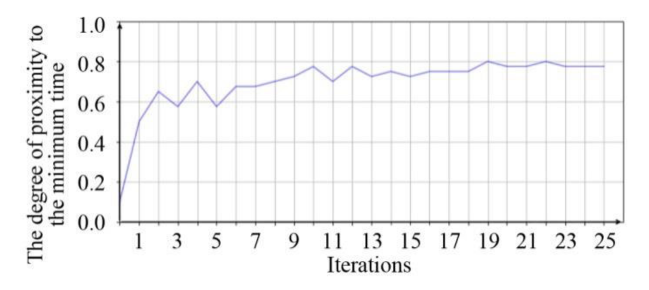

Fig. 9. The empirical LC based on the experimental data [11].

stabilizes in the range [0.75, 0.8]. This indicates rapid skill acquisition at the beginning and reaching a stable performance level after the 10th iteration.

Now we approximate the data using models (4) and (7)–(9) by selecting the values of the parameters $q(0)$, w, and u for model (4): $q_0 = 0.136$, $w = 0.490$, and $u = 0.170$. The approximation results are demonstrated in Fig. 10.

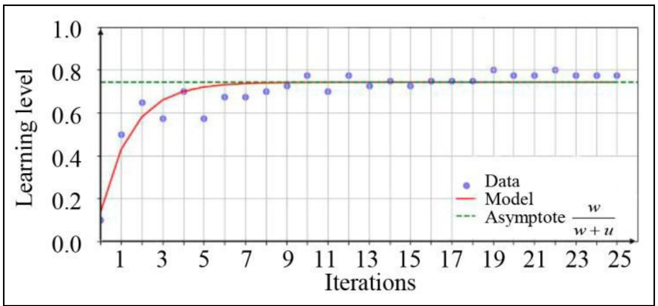

Fig. 10. The approximation of the experimental data [11] using model (4).

Model (4) describes a sharp increase in the learning level up to the 6th–7th iteration and an approach to the asymptote at a level of 0.75. The subsequent iterations do not lead to significant growth, and the factual values fluctuate around the model curve. This confirms that model (4) is capable of reflecting the process of rapid initial learning and further saturation.

The histogram of the deviations of model (4) can be found in Fig. 11.

The distribution is asymmetrical: most of the positive deviations are concentrated on the right-hand side (from 0.02 to 0.06), while the negative ones are more scattered. This may indicate a slight systematic error of the learning level on late iterations: the model "underestimates" the results compared to the factual plateau. Varying the value of τ_{min} preserves these regularities: the model's predictions are lower than the factual values, and the distribution of the deviations (see Fig. 11) has the same shape.

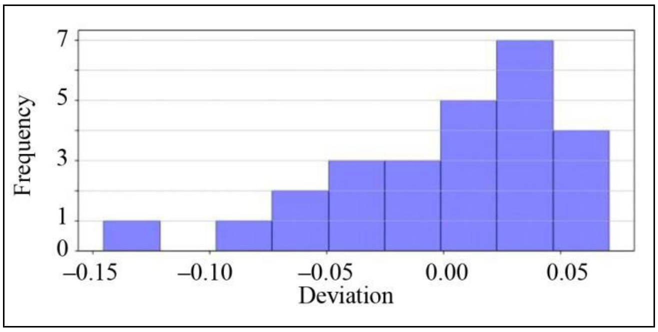

Fig. 11. The histogram of deviations of model (4) from the experimental data [11].

The Shapiro–Wilk statistic value is 0.929, and the *p*-value is 0.075. The normality hypothesis for the distribution is not rejected with a significance level of 0.05. (This statistical significance will be used below by default.) Similarly, the Epps–Pulley and Kolmogorov–Smirnov tests give no grounds for rejecting the hypothesis. The parameters of the distribution are $\mu \approx 0$ and $\sigma \approx 0.05$.

For the sake of comparison, we approximate the same data using model (7)–(9), in which the probability of mastering decreases over time. The resulting LC is demonstrated in Fig. 12.

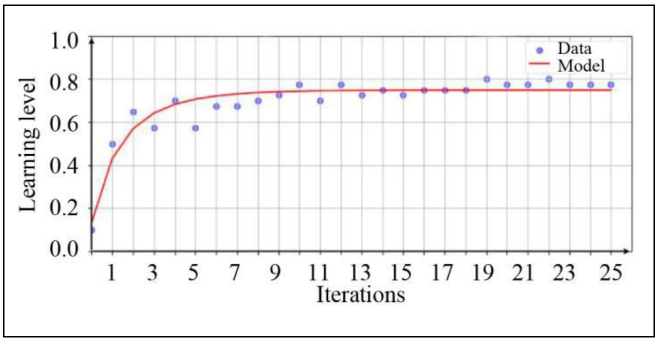

Fig. 12. The approximation of the experimental data [11] using model (7)–(9).

The curve reproduces the rapid growth of the learning level in the first 5–7 iterations and its saturation near 0.75. Unlike model (4), this approximation better reflects the smoothed dynamics of plateauing due to the time-varying probability of mastering.

The deviations from the empirical points are smaller, especially in the middle and late phases, indicating greater flexibility of the model in the case of a gradual decrease in the rate of mastering.

Numerical estimation yields the following values of the model parameters: $q_0 = 0.132$, $w_0 = 0.517$, and $\alpha = 0.417$. The histogram of the model deviations is provided in Fig. 13.

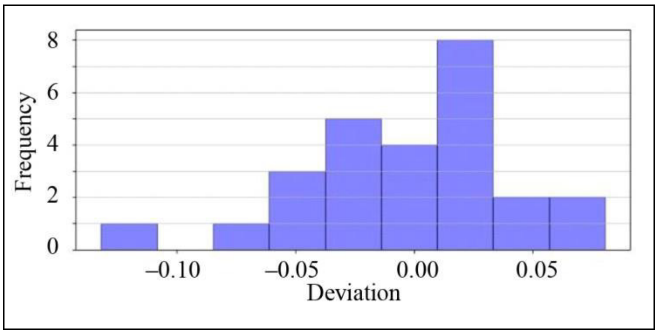

Fig. 13. The histogram of deviations of model (7)–(9) from the experimental data [11].

Most of the deviations are in the range from –0.05 to 0.05, with the maximum frequency occurring in regions close to zero. Unlike the previous model, the distribution here is more symmetrical.

The Shapiro–Wilk statistic value is 0.956, and the *p*-value is 0.311. The normality hypothesis for the distribution is not rejected with a significance level of 0.05. According to the Epps–Pulley and Kolmogorov–Smirnov tests, the normality hypothesis is not rejected as well. The parameters of the distribution are $\mu \approx 0$ and $\sigma \approx 0.05$. (In this example, the nonzero mean is more pronounced, i.e., already in the fourth decimal place: –0.000332 vs. 0.000000).

The data of the mental multiplication experiment reflect the process of gradually mastering a cognitive skill with a typical reduction in the task completion time. The empirical curve shows a rapid improvement phase followed by saturation, which corresponds to the classical shape of the iterative LC. The models from part I of the study [3] match the obtained data after their normalization by the transformation (10). Based on the analysis of deviations, there are no grounds to reject the normality hypothesis for their distribution with a mean close to zero. Therefore, the learning models from [3] are applicable to describe the formation of cognitive skills with regular and routine training.

3.3. The Visual-Motor Adaptation Experiment

Consider the data of the experiment from the paper [13]. The experiment consisted of four parts, each with three blocks of 20 trials. In each block, participants performed hand movements toward a target displayed on a screen. The key element of the experiment was the manipulation of the visual feedback of hand movement and the target's position. In the first version of the experiment, the visual feedback in the basic block was rotated 10º clockwise relative to the real hand movement. In the interference block, the feedback was rotated another 20º counterclockwise relative to the basic condition, while the target's position on the screen remained unchanged. The third block repeated the basic condition. The error value was recorded. Figure 14 shows the error value (averaged across all participants) as a function of the number of exercises.

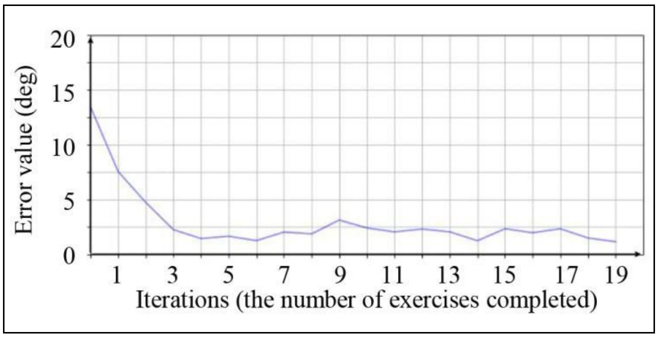

Fig. 14. The experimental data [13].

By the fifth iteration, the error decreases from approximately 15 to 2º and then stabilizes. The subsequent fluctuations range from 1.5 to 3º and demonstrate no clear trend. Consequently, the necessary motor pattern of actions is rapidly mastered at the beginning of the experiment, and a stable adaptation level is achieved with the minimum error.

Let us apply a transformation similar to formulas (1) and (10). For 0º and 20º as the minimum and maximum errors, respectively, we construct the LC based on the obtained data (Fig. 15).

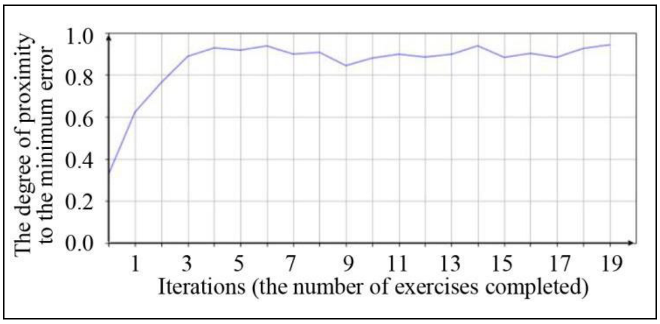

Fig. 15. The empirical LC based on the experimental data [13].

In the first three iterations, the curve rises from 0.4 to 0.85; then it reaches a level above 0.9, remaining there with slight fluctuations until the end of the experiment. The empirical curve demonstrates a smooth transition to the asymptote without significant dips; therefore, it can be described by standard saturation models.

We approximate the data using models (4) and (7)–(9). The result for model (4) is shown in Fig. 16.

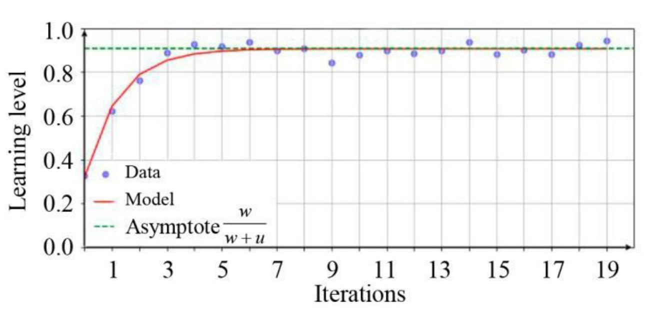

Fig. 16. The approximation of the experimental data [13] using model (4).

The LC for model (7)–(9) with the parameter values $q_0 = 0.319$, $w = 0.729$, and $u = 0.075$ reproduces a sharp increase in the first 3–5 iterations and then stabilizes at a level of about 0.9 (the model asymptote). The empirical data lie close to the model curve throughout the entire interval, but the deviations are observed in the central part (iterations 6–12): the model overestimates the predictions. Thus, despite the overall correspondence, model (4) may not fully reflect the peculiarities of the learning trajectory in the middle and late phases.

The histogram of the deviations between the predictions of model (4) and the experimental data can be found in Fig. 17.

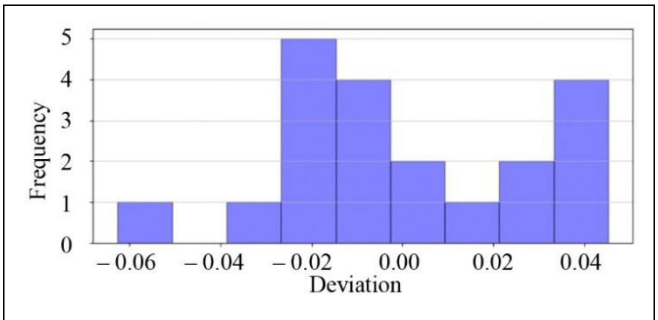

Fig. 17. The histogram of deviations of model (4) from the experimental data [13].

The distribution is asymmetrical: note a bias to the left, as most of the deviations lie in the negative region (from –0.03 to –0.01), and a pronounced peak on the right-hand side at a value of 0.04. This observation may indicate that the model underestimates the learning level in one region, overestimating it in another.

The Shapiro–Wilk statistic value is 0.943. There are no grounds to reject the normality hypothesis for the distribution of the deviations at a significance level of 0.05. The Epps–Pulley and Kolmogorov–Smirnov tests also give no reason to reject this hypothesis. The parameters of the distribution are $\mu \approx 0$ and $\sigma \approx 0.028$.

To refine the approximation, we employ model (7)–(9), in which the probability of mastering decreases over time. With this feature, it is possible to consider the effect of fatigue on the LC shape. The corresponding approximation of the experimental data is presented in Fig. 18.

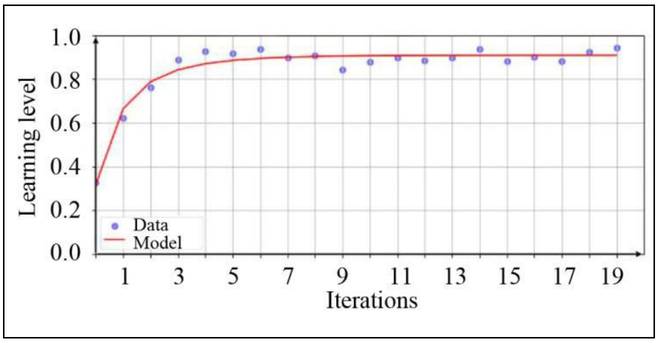

Fig. 18. The approximation of the experimental data [13] using model (7)–(9).

The model reproduces a sharp increase in the learning level in the first three iterations (from 0.2 to 0.85) and then accurately reaches a saturation level of about 0.93. Unlike model (4), here the curve describes the process more smoothly: the slope decreases gradually rather than abruptly. Numerical estimation yields the following values of the model parameters: $q_0 = 0.317$, $w_0 = 0.876$, and $\alpha = 0.432$. Figure 19 shows the histogram of the deviations between the predictions of model (7)–(9) and the experimental data.

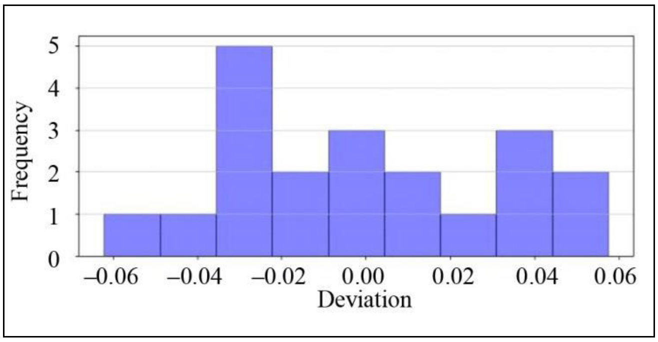

Fig. 19. The histogram of deviations of model (7)–(9) from the experimental data [13].

Most of the deviations are concentrated in the range from –0.03 to 0.05. Visually, the distribution looks more symmetrical compared to a similar histogram (Fig. 17) for model (4), but there is a slight bias to the negative side. This may indicate a partial overestimation of the model predictions at certain phases.

The Shapiro–Wilk statistic value is 0.968. The normality hypothesis of the distribution is not rejected with a significance level of 0.05. The Epps–Pulley and Kolmogorov–Smirnov tests also give no grounds for rejecting this hypothesis. The parameters of the distribution are $\mu \approx 0$ (but the mean has a nonzero value of 0.000267) and $\sigma \approx 0.032$.

The visual-motor adaptation experiment shows the rapid mastery of motor patterns when changing the visual feedback, with subsequent maintenance of the achieved level. The empirical curve is characterized by a sharp increase in the learning level and a smooth transition to saturation. The distributions of the deviations of the models under consideration are close to normal and have a nearly zero mean, which confirms the applicability of exponential laws to describe this type of motor tasks.

3.4. Breaks in the Ball-Tossing Experiment

The original experiment in [22] consisted of several periods; the first period has been discussed in subsection 3.1. The periods of active experimentation were interspersed with breaks: 22 months between the first and second stages, and 6 months between the second and third stages. Applying the transformation (1), we describe the data using models (3), (4), and (6)–(11) during the learning periods. The process of forgetting during the rest periods can be described by Model 2:

$$q(t) = q(0)e^{-ut}, \qquad (11)$$

where $q(0)$ is the learning level achieved by the end of the active experimentation period [3].

In the latter case, we calculate the probability of forgetting using the known values of the learning levels at the beginning and end of the break. The corresponding results are presented in Figs. 20–25.

In particular, Fig. 20 shows three periods of the experiment with long breaks between them. For each period separately, an approximation is constructed using model (4), and the period of forgetting between them is described by an exponential decrease in the learning level according to model (11). Clearly, after each rest period (between 100 and 657 days, and between 675 and 881 days), the level of mastering decreases, followed by a new growth.

To track the learning level dynamics, points A_1–A_3 and B_1–B_3 in Fig. 20 mark the starting and end points, respectively, the three active learning periods. The drop from B_1 to A_2, and then from B_2 to A_3, shows the effect of forgetting after long breaks. At the same time, the decrease between B_2 and A_3 is less pronounced, which agrees with the shorter second break compared to the first (6 vs. 22 months). The value at A_2 is not smaller than at A_1, and the value at A_3 is not smaller than at A_2; in other words, complete forgetting does not occur during the break. Also, the values at B_2 and B_3 are greater than at B_1, which may indicate the gradual accumulation of a stable skill or the formation of a more effective strategy during re-learning.

In general, with several periods of active learning or productive activity interspersed with breaks, the values in the sequences A_1–A_2–… and/or B_1–B_2–… may demonstrate the formation of a stable skill, complete forgetting, the effectiveness of subsequent repetitions, etc.

Consider an approximation approach where the data for each period are processed independently, without the intervals between them, since the data for assessing gradual forgetting were absent in the paper [22]. Figure 21 shows the approximation of each section using model (4) with the parameters selected separately for each active skill acquisition stage.

According to Fig. 21, the shape of the LC differs between periods: a smooth saturation in the first, a rapid growth, followed by a sharp plateau, in the second, and an almost linear curve in the third. This may indicate different regularities for the initial and repeated acquisition (consolidation) of skills.

The estimates of the curve parameters are: $q_0 = 0.188$, $w = 0.022$, and $u = 0.025$ (the first section); $q_0 = 0.186$, $w = 0.079$, and $u = 0.059$ (the second section); and $q_0 = 0.301$, $w = 0.070$, and $u = 0.049$ (the third section). The value of the parameter u during breaks is estimated as 0.00166 and 0.00281, respectively.

The histograms of the deviations between the predictions of model (4) and the experimental data in different periods are shown in Fig. 22.

In the first period, the distribution is close to Gaussian and centered near zero, which means a good approximation. In the second and third periods, the samples are small, but the deviations have no systematic bias. According to the Shapiro–Wilk, Epps–Pulley, and Kolmogorov–Smirnov tests, the normality hypothesis is not rejected for the deviations of model (4) from the experimental data.

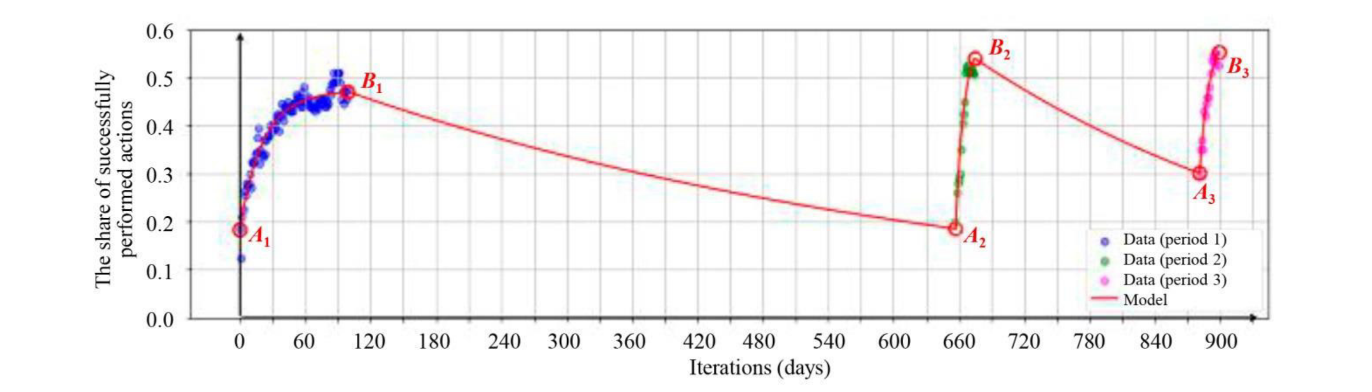

Fig. 20. The approximation of the experimental data [22] with breaks using models (4) and (11).

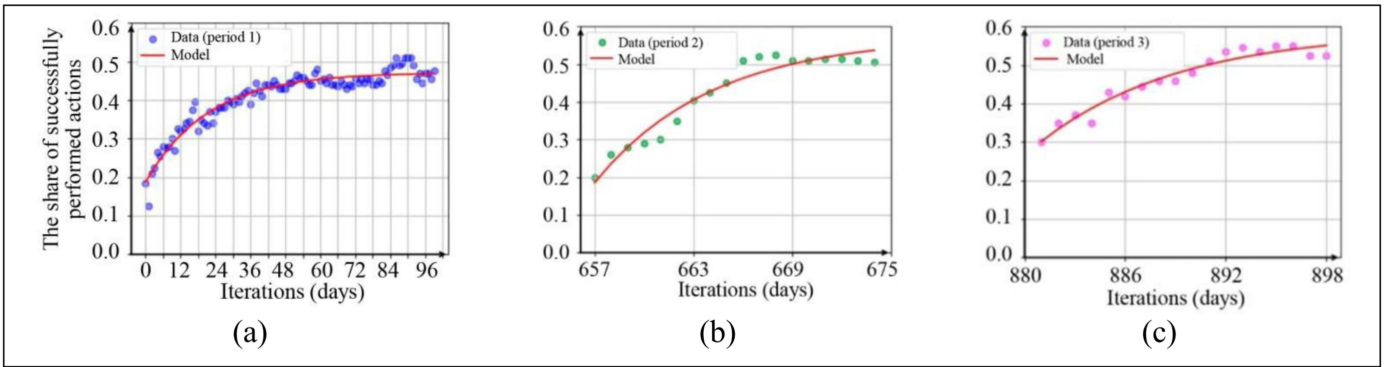

Fig. 21. The approximation of the experimental data [22] without breaks using model (4): (a) period 1, (b) period 2, and (c) period 3.

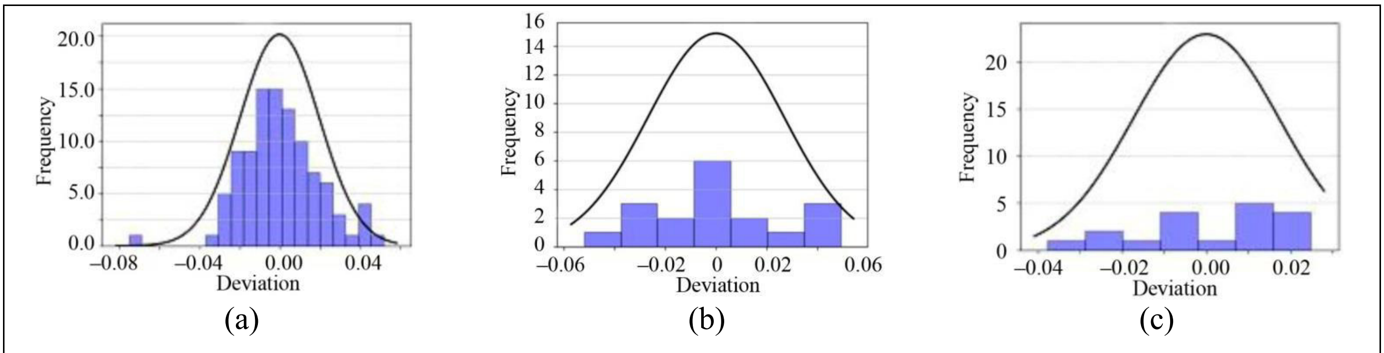

Fig. 22. The histogram of deviations of model (4) from the experimental data [22] without breaks: (a) period 1, (b) period 2, and (c) period 3.

To improve the approximation accuracy, we apply model (7)–(9), in which the probability of mastering a skill decreases over time. (Therefore, the effect of fatigue can be considered appropriately.) The corresponding results are presented in Fig. 23.

The model with fatigue more flexibly describes skill growth within the training period and, moreover, the decline in the rate of mastering over time. Unlike the model with constant parameters, where skill growth attenuates when approaching the asymptote, the model with the decreasing probability of mastery describes the slowdown due to the exhaustion of the mastery resource. This is especially important for tasks with long training periods, where the rate of skill growth may decrease due to the effect of fatigue rather than reaching the limit. In addition, plateauing is modeled not as "stopping" at a predetermined level but as continuing growth with a decreasing rate. This makes the approximation more realistic in the long term.

Due to the long intervals between the phases and no forgetting data during these periods, the constructed curve is not very informative in terms of dynamics. Therefore, further analysis employs only the periods of active skill acquisition (Fig. 24). The model reflects growth within each period, as well as the transition to a plateau and decelerated progress.

According to Fig. 24, similar to model (4), the LC shape changes depending on the period: smooth saturation in the first, and rapid growth and plateauing in the others.

The estimates of the curve parameters are: q_0 = 0.183, w_0 = 0.018, and α = 0.041 (the first section); q_0 = 0.185, w_0 = 0.069, and α = 0.099 (the second section); and q_0 = 0.301, w_0 = 0.051, and α = 0.089 (the third section). The value of the parameter u during breaks is estimated as 0.00167 and 0.00284, respectively.

The histograms of the model deviations can be found in Fig. 25.

According to the deviation histograms, model (7)–(9) generally provides a more symmetrical distribution of errors compared to model (4) (see Fig. 22). In the

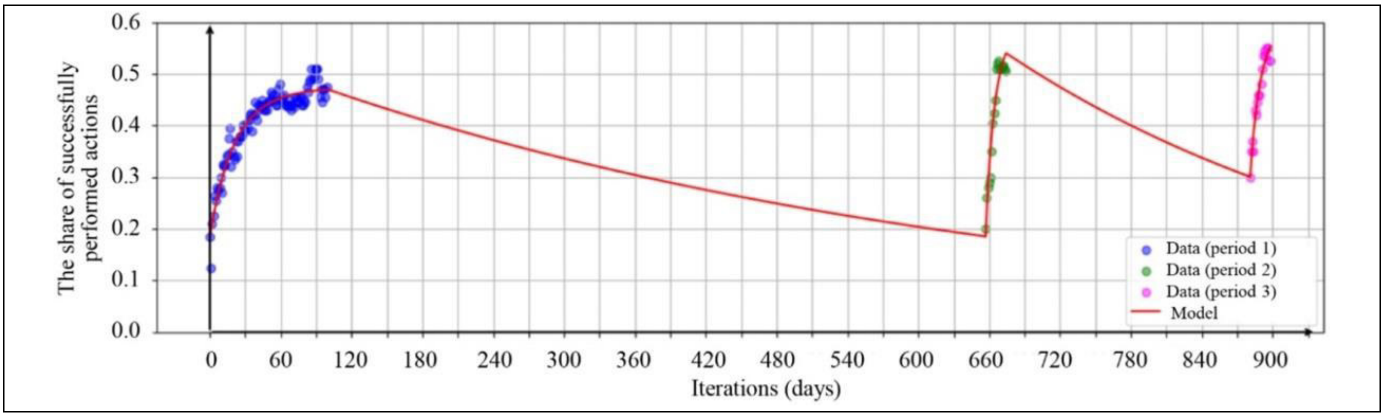

Fig. 23. The approximation of the experimental data [22] with breaks using models (7)–(9) and (11).

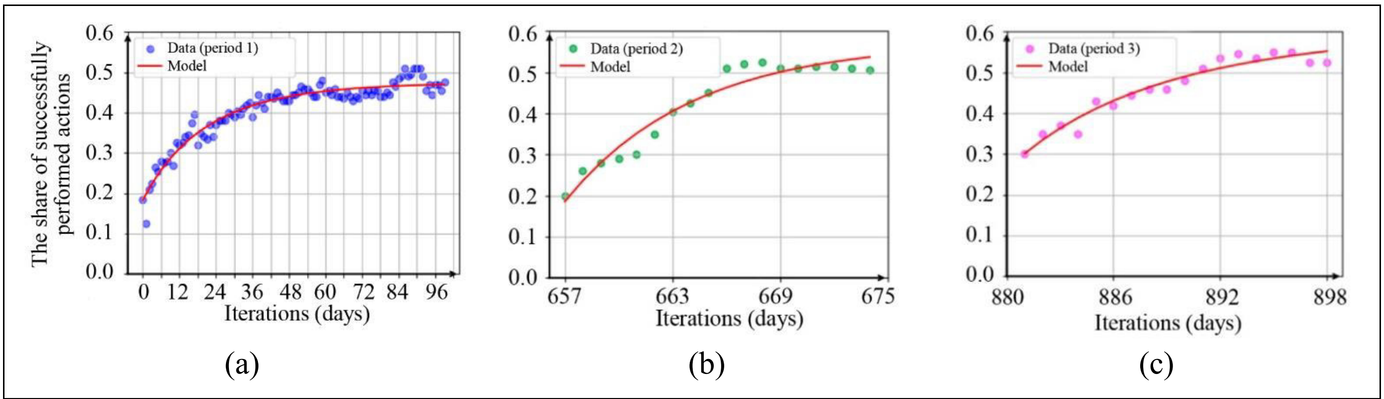

Fig. 24. The approximation of the experimental data [22] with breaks using models (7)–(9): (a) period 1, (b) period 2, and (c) period 3.

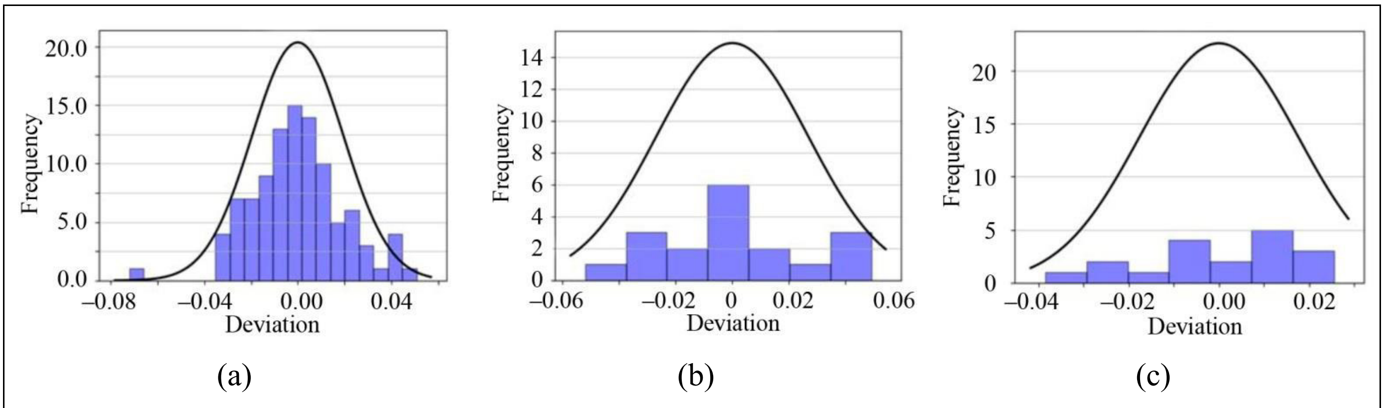

Fig. 25. The histogram of deviations of models (7)–(9) from the experimental data [22] without breaks: (a) period 1, (b) period 2, and (c) period 3.

first and third periods, the deviations are grouped closer to zero, and their amplitude is lower (a more accurate approximation). In the second period, there is a slight asymmetry, albeit less pronounced than under the approximation by the constant-parameter model. Thus, the model with a decreasing probability of mastering better considers the characteristics of the learning rate in different periods.

Based on the Shapiro–Wilk, Epps–Pulley, and Kolmogorov–Smirnov tests, the normality hypothesis of the deviations of models (4) and (7)–(9) from the factual data is not rejected. However, the parameters have to be selected separately for each period: the model with the parameter values obtained in the first stage does not match the data in the subsequent stages.

In addition, it is impossible to analyze the deviations of the model from the experimental data during the rest stages for the values from the paper [22]: the experiment does not include measurements during breaks. Moreover, such measurements would be a training themselves (i.e., part of the learning process) and could affect the final result.

Note also that with such a long break between the experiment stages, it is necessary to select "new" model parameters, since the parameter estimates obtained in the previous stages do not match the available data. The forgetting parameter during rest is not comparable to its analog during training, partly due to the significant difference in the length of the stages.

The ball-tossing experiment with long breaks shows the differences between the initial and repeated acquisition of the skill. According to the step-by-step approximation of empirical data, the models from part I of the study [3] describe the learning level dynamics, and the deviations from the empirical data are normally distributed with a nearly zero mean. There are several controllable factors in the design of experiments with breaks. It is generally possible to control the duration of training periods and breaks in order to minimize skill loss due to forgetting and, at the same time, provide an opportunity to "restart" training. This can relieve systematic fatigue and lead to new strategies for achieving the learner's goals.

4. A COMPARISON OF MODELS BASED ON THE RESULTS OBTAINED

The approximation quality of Models 1, 3, 4, and 5 from part I of the study [3] with constant and decreasing probability of mastering (without and with fatigue, respectively) is compared in the table below. For each approximation, the model parameters were selected using the least squares method. Next, the coefficient of determination R^2 was calculated to assess the model's correspondence to the experimental data.

According to the comparison results, in most cases, the model with fatigue or forgetting better matches the data compared to the ones without these processes. In the absence of fatigue or forgetting, the model reproduces the general slowly asymptotic nature of the curve, but worse reflects a gradual decline in the learning rate in the later stages.

At the same time, the difference in the values of the coefficient of determination R^2 between models is often small; see Models 4 and 5. A particularly noticeable improvement is observed in the description of cognitive skills (mental multiplication), which may be associated with a more pronounced decline in the rate of mastering.

The exception is the visual-motor adaptation experiment, where the constant-parameter model has a slightly higher coefficient of determination. This may be due to the task specifics: unlike the other experiments, here the participant was clearly given a final goal (hit the right spot), and it was achieved via constant feedback. Such experiments reduce cognitive load: there is neither constant strategy revision nor striving for an abstract maximum. As a result, fatigue either disappears or has little effect on the skill acquisition dynamics. In other tasks, participants acted without a clear endpoint, which increases uncertainty, complicates focus, and may contribute to fatigue.

In the final analysis, the models with three or more parameters (Models 3–5) demonstrate significantly greater accuracy compared to the two-parameter counterpart (Model 1). In two out of the three cases, the model with a decreasing probability of mastering (Model 4) better approximates the data compared to the constant-parameter one (Model 3). The advantage of Model 5 over Model 4 is minimal: taking fatigue into account mainly contributes to accuracy improvement, whereas the additional consideration of forgetting during training does not significantly affect the value of R^2 in the experiments presented.

Based on these outcomes, fatigue consideration seems reasonable when modeling tasks with prolonged load or a pronounced decrease in the rate of mastering; for short and intensive experiments, constant-parameter models may be sufficient.

The quality of models: a comparison

Experiment	Model			
	Model 1 w = const	Model 3 w = const, u = const	Model 4 $w = w(t) = w_0 e^{-\alpha t}$	Model 5 $w(t) = w_0 e^{-\alpha t}$, u = const
Ball tossing (stage 1)	0.76076	0.93025	0.93200	0.93473
Mental multiplication	0.56702	0.87209	0.89110	0.94287
Visual-motor adaptation	0.65981	0.96134	0.94981	0.96137

CONCLUSIONS

According to the review of experimental datasets on the acquisition of motor and cognitive skills, the processes of mastering, forgetting, fatigue, and rest can manifest themselves in different forms depending on the experiment type. Mastering is present in all datasets, which confirms the applicability of general iterative learning models. Forgetting and fatigue manifest themselves in different ways, i.e., through an increase in errors and execution time or a decrease in accuracy after pauses (rest). The parameters (time, the history of actions, the number of components, and fatigue rate) serve to tune the model to a particular experiment; the resulting model can be used for comparison, prediction, and control of the learning process.

The unified experimental data processing principles (see the discussion above) demonstrate that various criteria for the learning level (temporal, speed- and accuracy-related, and informational) can be reduced to the accuracy-related criterion under which the LC has a slowly asymptotic nature. In addition, absolute units of measurement of the learning level can be converted into relative (probabilistic) ones; therefore, general mathematical models can be used to describe the experience acquisition process based on a wider range of experiments.

The above examples of processing experimental results show that the learning models match the available data—the statistical tests give no grounds to reject the hypothesis that the model deviations are independent and identically distributed realizations of a Gaussian random variable with zero mean.

According to the ball-tossing experiment with a break after a long interval, the initial learning level drops, but after several days of training, the result is higher than in the first period. This means a positive effect of the break on the terminal learning level; but it is impossible to determine its optimal duration to confirm the conclusions from part I of the study [3] based on the available evidence.

In the examples considered, the models can be used to recommend an appropriate organization of the learning process. For instance, when describing the experience acquisition processes in the mental multiplication and visual-motor adaptation experiments by model (4), the LC almost reaches the asymptote much earlier than the last iteration. Hence, given the experiment setup and the confirmed achievement of the maximum learning level, it is advisable to terminate the learning process or change the learning strategy.

By considering fatigue, one can describe the decrease in the skill acquisition rate during training: the model shows a gradual transition from fast progress to a slowdown, without interrupting growth at a fixed level. This approach to learning modeling may be particularly important for long-term skill acquisition. For three of the four experiments presented, the coefficient of determination of the model with a decreasing probability of mastery has been higher. A noticeable improvement has been observed in the description of the cognitive skill (mental multiplication), where the acquisition rate slows down the most. However, in the visual-motor adaptation experiment, model (4) has shown the best value of the coefficient of determination R^2.

Open datasets for building and testing learning models

Dataset or publication	Description	The field of activity	Data	Components of learning		Reference
				Processes	Parameters	
ASSISTments	Contains data on solving math tasks. One question may have several skill tags. Hints are provided in the tasks.	Education. Solving math tasks	User tag, task tag, correct answer, user answer, the number of attempts, the number of hints used, and skills acquired. The number of participants: over 60 000 persons.	**Mastering:** a decrease in the use of hints for tasks within a skill; an increase in the number of correct answers to tasks within a skill. **Forgetting:** an increase in the use of hints for tasks within a skill; a decrease in the number of correct answers to tasks within a skill. **Fatigue:** with the course of time, more hints are used, and the number of incorrect answers to tasks within different skills increases. **Rest:** a break in interaction with the system between sessions on the same day.	**Time:** time stamps. **History:** the correctness of task completion. **The number of skills acquired:** 1 or more, training to complete tasks for different skills.	2009 data: https://sites.google.com/site/assistmentsdata/home/2009-2010-assistment-data/skill-builder-data-2009-2010 (Accessed April 28, 2025.) 2012 data: https://sites.google.com/site/assistmentsdata/datasets/2012-13-school-data-with-affect (Accessed April 28, 2025.) 2015 data: https://drive.google.com/file/d/0B_hO8cnpcIMgUGZzRnh3bHJrSjQ/view?resourcekey=0-dGtan-IMFc3IjQ749-FgQA (Accessed April 28, 2025.)
Duolingo Spaced Repetition Data	Contains data on memorizing words and their use in different contexts on the Duolingo platform.	Education. Learning of foreign languages	User tag, language learned, lexeme tag, the number of lexeme demonstrations to the user before the lesson, the number of correct word definitions before the lesson, the number of lexeme demonstrations during the lesson, and the number of correct word definitions during the lesson. The number of participants: over 150 000 persons.	**Mastering:** maintaining the accuracy of answers to word translations over time. **Forgetting:** an incorrect answer after a series of correct answers to word translations. **Fatigue:** an increase in the number of incorrect answers to all word translations. **Rest:** a break in interaction with the application on the same day. **Selection of the skill acquired:** by the system. For example, the system more frequently generates words that are commonly forgotten by the user.	**Time:** time stamps. **History:** the correctness of task completion. **The number of skills acquired:** 1 or more, training to translate different words.	https://www.kaggle.com/datasets/aravinii/duolingo-spaced-repetition-data (Accessed February 7, 2025.)

Table (continued)

Dataset or publication	Description	The field of activity	Data	Components of learning		Reference
				Processes	Parameters	
Junyi Academy Online Learning Activity Dataset	Contains data on user activity on the platform designed for solving tasks in various disciplines.	Education	Three tables with metadata of students and exercises, attempts to solve tasks, time stamps, and the number of hints used. The number of participants: over 72 000 persons.	**Mastering:** a decrease in errors in repeated attempts to solve a task. **Forgetting:** return to errors after a progression of correct answers to the task. **Fatigue:** with the course of time, more hints are used, and the number of incorrect answers to different tasks increases. **Rest:** a break in interaction with the system between sessions on the same day.	**Time:** time stamps. **History:** the history of answers and interactions. **The number of activity components:** 1 or more, training to complete different tasks.	https://www.kaggle.com/datasets/junyiacademy/learning-activity-public-dataset-by-junyi-academy/data?select=Log_Problem.csv (Accessed February 24, 2025.)
Junyi Academy Math Practicing Log	Contains data on user activity on the platform when solving math tasks.	Education. Solving math tasks	Metadata of students and exercises, attempts to solve tasks, time stamps, and the number of hints used. The number of participants: over 2000 persons.	**Mastering:** a decrease in errors when retrying to solve a task. **Forgetting:** return to errors after a progression of correct answers to the task. **Fatigue:** with the course of time, more hints are used, and the number of incorrect answers to different tasks increases. **Rest:** a break in interaction with the system between sessions on the same day.	**Time:** time stamps. **History:** the history of answers and interactions. **The number of skills acquired:** 1 or more, training to complete different tasks.	https://pslcdatashop.web.cmu.edu/DatasetInfo?datasetId=1275 (Accessed February 24, 2025.)
EdNet	Contains user data on learning on the multi-profile platform. Includes answers to questions and interactions with the system.	Education	Four types of datasets with user tags, time stamps, question tags and correct answers, user answers, as well as user actions on the platform, such as viewing video lectures and materials. The number of participants: over 780 000 persons.	**Mastering:** an increase in the number of correct answers to the task. **Forgetting:** a decrease in the accuracy of answers after a progression of correct answers within the task. **Fatigue:** a decrease in activity or an increase in errors on different tasks. **Rest:** periods of inactivity between sessions during the day.	**Time:** time stamps. **History:** the history of answers and interactions. **The number of skills acquired:** 1 or more, training to complete different tasks.	https://github.com/riiid/ednet (Accessed May 26, 2025.)

Table (continued)

Dataset or publication	Description	The field of activity	Data	Components of learning		Reference
				Processes	Parameters	
Codeforces Users Submissions	Contains data on user activity on the platform designed for solving programming tasks.	Education. Solving programming tasks	User tag, task tag, task difficulty, user answer, and response time. The number of participants: over 15 000 persons.	**Mastering:** an increase in successful task solutions. **Forgetting:** a decrease in answer accuracy after a progression of correct answers within the task. **Fatigue:** an increase in the time taken to solve tasks within the competition.	**Time:** time stamps. **History:** the complexity of previously solved tasks. **The number of skills acquired:** 1 or more, training to complete different tasks.	https://huggingface.co/datasets/denkCF/UsersCodeforcesSubmissionsEnd2024 (Accessed May 20, 2025.)
A Learning Curve Equation as Fitted to Learning Records [11]	Contains data on solving a set of 63 exercises (the mental multiplication of two two-digit numbers) by a participant.	Education. Solving math tasks	Module completion time. The number of participants: 1 person.	**Mastering:** a decrease in the time spent on completing the exercises. **Forgetting:** an increase in the time spent on completing the exercises.	**Time:** iterations. **History:** the results of previous days. **The number of skills acquired:** 1, the same skill is trained.	https://www.semanticscholar.org/paper/A-learning-curve-equation-as-fitted-to-learning-Barlow/1eb32cea20d374129b85069c93feeeb75330882c
Injury Prediction in Competitive Runners	Contains a detailed training log of high-level runners over seven years (2012–2019). Includes data on runners competing in distances from 800 m to the marathon.	Sports	Runner number, running distance, the number of trials, the presence of strength training, the hours of alternative activities, and the subjective assessments of the runner's condition before and after training sessions. The number of participants: 74 persons.	**Fatigue:** runners' subjective assessments of fatigue after training sessions, the influence of the number of different training sessions and distances.	**Time:** dates of training sessions. **History:** previous training sessions. **The number of skills acquired:** 1 or more, training to run different distances.	https://dataverse.nl/dataset.xhtml?persistentId=doi:10.34894/UWU9PV (Accessed February 28, 2025.)
An Extensive Experiment in Motor Learning and Re-Learning [22]	Contains data on attempts to throw a ball into a box. The ultimate objective of the experiment was to study the learning capability by trial and error when	Movement	The number of balls scored. The number of participants: 1 person.	**Mastering:** an increase in the number of balls scored. **Forgetting:** a decrease in the number of balls scored, especially after a break. **Rest:** breaks in the experiment.	**Time:** iterations by day. **History:** the results of previous days.	https://www.semanticscholar.org/paper/An-Extensive-Experiment-in-Motor-Learning-and-Braden/0f288c34ce089188926abfdbe87e40ba7399b499

Table (continued)

Dataset or publication	Description	The field of activity	Data	Components of learning		Reference
				Processes	Parameters	
	learning is spread out over time, and especially to study the re-learning curve.				**The number of skills acquired:** 1, the same skill is trained.	
GradualTwoRate	Contains data on the adaptation of different-age persons to sudden and gradual disturbances in the external environment. The experiment was to train a person to coordinate his/her movements in accordance with visual perception, which was changed using a special installation.	Movement. Adaptation to changes in virtual feedback	Participant number, iteration number, rotation angle, and deviations from the trajectory. The number of participants: 136 persons.	**Mastering:** adaptation to disturbances. **Fatigue:** gradual deterioration of results compared to previous iterations.	**Time:** iterations. **History:** the angles and trajectories in iterations prior to the current one. **The number of skills acquired:** 1 or more, different trajectories are trained.	https://osf.io/c5ezv/ (Accessed March 13, 2025.)
Motor Learning Experiments	Contains data on person's adaptation to environmental disturbances. The experiment was to train a person to coordinate his/her movements in accordance with visual perception, which was changed using a special installation.	Movement. Adaptation to changes in virtual feedback	Participant number, attempt number, target angle, achieved angle, and deviations from the trajectory. The number of participants: 10 persons.	**Mastering:** adaptation to disturbances at a certain angle. **Forgetting:** deterioration of results in attempts after a break. **Fatigue:** gradual deterioration of results compared to previous iterations.	**Time:** iterations. **History:** the angles and trajectories in iterations prior to the current one. **The number of skills acquired:** 1 or more, different trajectories are trained.	https://figshare.com/articles/dataset/Data_from_motor_learning_experiments/957526/1 (Accessed March 13, 2025.)

REFERENCES

1. Belov, M.V. and Novikov, D.A., Models of Experience, *Control Sciences*, 2021, no. 1, pp. 37–52. DOI: 10.25728/cs.2021.1.5.
2. Belov, M.V. and Novikov, D.A., *Models of Technologies*, Cham: Springer, 2020.
3. Grebenkov, D.I., Kozlova, A.A., Lemtyuzhnikova, D.V., and Novikov, D.A., Models of Fatigue and Rest in Learning. Part I: Extension of the General Iterative Learning Model, *Control Sciences*, 2025, no. 3, pp. 24–31.
4. Novikov, D.A., *Zakonomernosti iterativnogo naucheniya* (Regularities of Iterative Learning), Moscow: Institute of Control Sciences RAS, 1998. (In Russian.)
5. Novikov, A.M., *Protsess i metody formirovaniya trudovykh umenii* (The Process and Methods to Form Labor Skills), Moscow: Vysshaya Shkola, 1986. (In Russian.)
6. *Duolingo Spaced Repetition Data*. URL: https://www.kaggle.com/datasets/aravinii/duolingo-spaced-repetition-data (Accessed February 7, 2025.)
7. *Assistments Data 2009-2010*. URL: https://sites.google.com/site/assistmentsdata/home/2009-2010-assistment-data/skill-builder-data-2009-2010 (Accessed April 28, 2025.)
8. *Data from Motor Learning Experiments*. URL: https://figshare.com/articles/dataset/Data_from_motor_learning_experiments/957526/1 (Accessed March 13, 2025.)
9. *Replication Data for: Injury Prediction in Competitive Runners with Machine Learning*. URL: https://dataverse.nl/dataset.xhtml?persistentId=doi:10.34894/UWU9PV (Accessed February 28, 2025.)
10. *Junyi Academy Online Learning Activity Dataset*. URL: https://www.kaggle.com/junyiacademy/learning-activity-public-dataset-by-junyi-academy (Accessed February 4, 2025.)
11. Barlow, M., A Learning Curve Equation as Fitted to Learning Records, *Psychological Review*, 1928, vol. 35, no. 2, pp. 142–160.
12. *Codeforces Users Submissions Dataset (End of 2024)*. URL: https://huggingface.co/datasets/denkCF/UsersCodeforcesSubmissionsEnd2024 (Accessed May 13, 2025.)
13. *GradualTwoRate Dataset*. URL: https://osf.io/c5ezv/ (Accessed March 13, 2025.)
14. *IS0 5479: Statistical Interpretation of Data - Tests for Departure from the Normal Distribution*, 1997.
15. Goryainov, V.B., Pavlov, I.V., Tsvetkov, G.M., et al., *Matematicheskaya statistika* (Mathematical Statistics), Zarubin, V.S. and Krishchenko, A.P., Eds., Moscow: Bauman Moscow State Technical University, 2001. (In Russian.)
16. Nazarov, A.A., *Statistical-methods*. URL: https://github.com/AANazarov/Statistical-methods/tree/master/Epps-Pally%20test (Accessed February 28, 2025.)
17. Epps, T.W. and Pulley, L.B., A Test for Normality Based on the Empirical Characteristic Function, *Biometrika*, 1983, vol. 70, no. 3, pp. 723–726.
18. Shapiro, S. and Wilk, M., An Analysis of Variance Test for Normality (Complete Samples), *Biometrika*, 1965, vol. 52, no. 3-4, pp. 591–611.
19. Royston, P., Approximating the Shapiro–Wilk W-test for Non-normality, *Statistics and Computing*, 1992, vol. 2, pp. 117–119.
20. Bol'shev, L.N. and Smirnov, N.V., *Tablitsy matematicheskoi statistiki* (Tables of Mathematical Statistics), Moscow: Nauka, 1983. (In Russian.)
21. Lemeshko, B.Yu., *Kriterii proverki otkloneniya raspredeleniya ot normal'nogo zakona. Rukovodstvo po primeneniyu* (Criteria for Testing the Departure of a Distribution from the Normal Law. Application Guidelines), Novosibirsk: Novosibirsk State Technical University, 2014. (In Russian.)
22. Braden, S., An Extensive Experiment in Motor Learning and Re-Learning, *Journal of Educational Psychology*, 1924, vol. 15, no. 5, pp. 313–315.

This paper was recommended for publication
by A.A. Voronin, a member of the Editorial Board.

Received August 14, 2025,
and revised October 20, 2025.
Accepted October 20, 2025.

Author information

Grebenkov, Dmitry Igorevich. Mathematician, Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ grebenkov-d-i@mail.ru
ORCID iD: https://orcid.org/0009-0002-7085-5912

Kozlova, Anastasiia Andreevna. Engineer, Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ sankamoro@mail.ru
ORCID iD: https://orcid.org/0009-0005-6105-121X

Lemtyuzhnikova, Dar'ya Vladimirovna. Cand. Sci. (Phys.–Math.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ darabbt@gmail.com
ORCID iD: https://orcid.org/0000-0002-5311-5552

Novikov, Dmitry Aleksandrovich. Academician, Russian Academy of Sciences; Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia
✉ novikov@ipu.ru
ORCID iD: https://orcid.org/0000-0002-9314-3304

Cite this paper

Grebenkov, D.I., Kozlova, A.A., Lemtyuzhnikova, D.V., and Novikov, D.A., Models of Fatigue and Rest in Learning. Part II: Motor and Cognitive Skills. *Control Sciences* 6, 50–68 (2025).

Original Russian Text © Grebenkov, D.I., Kozlova, A.A., Lemtyuzhnikova, D.V., Novikov, D.A., 2025, published in *Problemy Upravleniya*, 2025, no. 6, pp. 58–79.

This paper is available under the Creative Commons Attribution 4.0 Worldwide License.

Translated into English by *Alexander Yu. Mazurov*,
Cand. Sci. (Phys.–Math.),
Trapeznikov Institute of Control Sciences,
Russian Academy of Sciences, Moscow, Russia
✉ alexander.mazurov08@gmail.com