The Problem of the Limitations of the Educational Model Experiment on Population Genetics and Its Solution

Elena V. Komarova[1][a] and Arnold E. Kiv[2][b]
[1]*Immanuel Kant Baltic Federal University, 14 Aleksandra Nevskogo Str., Kaliningrad, 236041, Russia*
[2]*Ben-Gurion University of the Negev, P.O.B. 653, Beer Sheva, 8410501, Israel*

Keywords: Population Genetics, Repeatability of the Experiment, Massiveness of the Experiment, Ideal Population, Hardy-Weinberg's Law.

Abstract: The difficulty of conducting an educational model experiment on population genetics is in meeting the requirements of mass and replication. The evolution of a model experiment to study the Hardy-Weinberg law according to the methods proposed by the authors is considered – from the use of material models and manual counting of alleles and genotypes to the transition to automatic random distribution of elements of the genotypic structure of the population and automatic calculation of the resulting indicators. The technique of fully automated modeling of the genetic structure of the population allows to increase the size of the model sample by orders of magnitude. When using the technique in group work of students, it becomes possible to demonstrate the essence and differences of technical and biological replication as requirements for organizing a biological experiment. The technique is currently developed to work only with very large ideal populations.

1 INTRODUCTION

The academic subject "Biology" is a didactically adapted system of scientific biological knowledge. In the natural sciences, and biology is undoubtedly one of them, experiment is one of the main methods of research (Nechypurenko et al., 2021). This is what makes it possible, on the basis of the various factual material obtained, to make broad generalizations, to proceed to the establishment of connections, patterns that allow deeper penetration into the essence of the phenomena under study. Much has already been said about the experiment in biological science, about its types, methods, requirements for organization, limitations and difficulties of application. A huge number of scientific works are devoted to the history of the experimental method in biology. However, we were and are interested at the moment in the experimental method from the point of view of the possibilities of its use in biological education. In addition, narrowing down the subject area of interest to us, it is worth noting that the model experiment occupies a special place in high school. This allows you to create models of real objects and prototype the processes that occur with them in reality. In previous works, the authors has covered some aspects of this issue (Komarova and Azaryan, 2018; Komarova and Starova, 2020).

Modern course of biology in high school is based on the fundamental theoretical generalizations of basic biological science – scientific theories and laws. Fundamental genetic laws, classically studied by high school students, are laws of heredity of Mendel. Given the trends of development of modern biological sciences, namely, the development of theoretical biology, the main issues which are problems of genetics, ecology, evolution, law of genetic equilibrium concentrations (the law of Hardy-Weinberg) is considered as a fundamental law, the disclosure of which to high school students is aimed at understanding by them of the mechanism of evolution in general. This law reveals the regularities of functioning of living at the population – species level, including time frames.

Students' mastering of the patterns of population genetics and associated evolutionary theory is one of the most complex issues in biology course in high school. Studies (Hammersmith and Mertens, 1990; Mertens, 1992; Moore, 1994; Maret and Rissing, 1998; Mukhopadhyay et al., 2014; Pongsophon et al., 2007) confirm this.

We have conducted a survey among 52 high school students to ascertain their level of knowledge

[a] https://orcid.org/0000-0002-3476-3351
[b] https://orcid.org/0000-0002-0991-2343

Komarova, E. and Kiv, A.
The Problem of the Limitations of the Educational Model Experiment on Population Genetics and Its Solution.
DOI: 10.5220/0010923700003364
In *Proceedings of the 1st Symposium on Advances in Educational Technology (AET 2020) - Volume 1*, pages 272-286
ISBN: 978-989-758-558-6

about the essence of law of genetic equilibrium concentrations, its value for the understanding of the factors and directions of the evolutionary process.

The tasks were as following:

1. Specify the mathematical equation of the law of Hardy–Weinberg (multiple answers are allowed):
   (a) $p+q=1$;
   (b) $(p+q)^2=1$;
   (c) $p^2+2pq+q^2=1$;
   (d) $p^2+pq+q^2=1$;
   (e) $p+2pq+q=1$.
2. Specify an equation describing the genotypic structure of the population (multiple answers are allowed): (see the answers to the assignment 1).
3. Specify the equation describing allelic population structure: (see answers to the assignment 1).
4. What are the conditions of validity of the law of equilibrium gene concentrations. Students were asked to solve three problems for the application of the law of equilibrium of gene concentrations.

The results of the survey are presented in figures 1–3.

Conditions of validity of the law, according to the student, were as following: population sizes are large – 29% of respondents, mating occurs at random – 24%, new mutations do not occur – 18%, all genotypes are equally fertile – 12%, generations do not overlap – 12%, there is no exchange of genes with other populations – 18%, the genes are in the autosomes and not in sex chromosomes - 18%, individuals of different genotypes are equally viable – 12%.

The obtained results allow to formulate the following conclusions: students insufficiently understood the description of the essence of the law of Hardy-Weinberg with two equations, namely, the definition of allelic and genotype structure of the population; the students are confused about variables included in the equations; knowledge about the conditions of the law is fragmentary. None of the respondents began to address two of the proposed problems, three of the respondents solved the third task incorrectly.

The results of the survey suggest the presence of formalism in high school students' knowledge about the law of genetic equilibrium concentrations. Formal approach to training lies in the mechanical memorization of educational material without enough understanding of its content. The low level of development of knowledge about the law of genetic equilibrium concentrations is one of the reasons for the difficulties of the students in understanding of the evolutionary content for the understanding of population genetics and population human genetics in particular.

Simulation, particularly computer simulation is one of the most effective training methods for demonstrating to students of the essence of complex biological processes, including genetic and evolutionary processes in natural populations.

In the process of studying the topic of the display of experimental method at the level of school biological education, transformation of ideas about how it is possible to implement the experimentation with complex biological systems, including those inaccessible to the student for direct study, we initially started out from the following. An educational biological experiment should maximally meet the requirements that are put forward for scientific biological experimentation. These, in particular, are the reliability in essence, the rule of single difference, replication, mass nature. From the mid-XXth century, a lot of attention has been paid to the organization of a school biological experiment, moreover to its various types, differing both in the object of research (botanical, zoological, physiological tests, functional tests, etc.), and in the form of carrying out under the conditions of school laboratory, class (demonstrational, laboratorial, mental) (Binas et al., 1990; Voronin and Mash, 1983; Shamrai and Zadorozhnyi, 2003; Frolov, 2007; Brunovt et al., 1973; Bazykin, 1988; Borodin, 1987; Bulaeva, 1977; Sidorova, 2009). Regardless of the type and form of carrying out, all various types of educational biological experiment must meet the abovementioned requirements in order that the results obtained were maximum consistent.

The biggest difficulties in the educational process are caused by the observance of such requirements as replication and mass nature. In other words, to ensure the veracity of results, the educational experiment should be conducted several times using a sufficiently large number of objects.

It is difficult to implement both the first and the second condition in the educational process due to the following reasons:

- firstly, the temporal limitations of the educational process;
- secondly, due to the inaccessibility of objects for study in the required quantity;
- thirdly, in the principle of inaccessibility of some objects and processes for direct study, primarily due to their objective specificity: either too small (organic molecules, cells, viral particles), or too large (populations).

Let's turn our attention to these reasons, possible ways of their elimination.

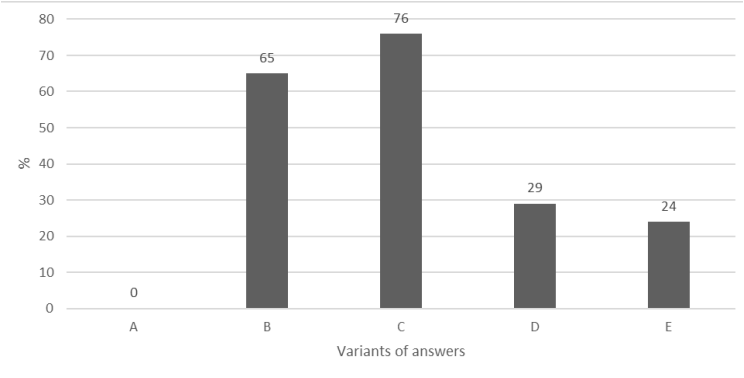


Figure 1: The results of the response to the first task.

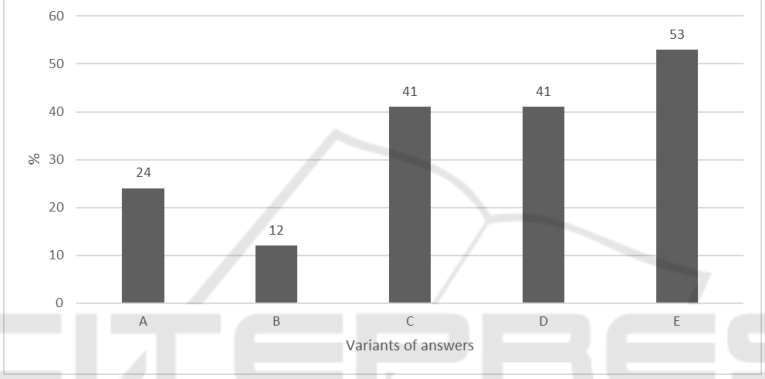


Figure 2: The results of the response to the second task.

Analysis of scientific literature in regards to the experimentation in biology showed that since the end of the 1980s, one of the actively discussed problems became the problem of pseudoreplication in ecological and biological research. In the classical variation, from the moment of publication of the first paper on this topic, pseudoreplication was considered as a negative experimental practice (Hurlbert, 1984). Even now, one of the criteria by which reviewers evaluate the submitted paper for a journal indexed in the authoritative international scientometric Scopus and Web of Science databases is the true and pseudoreplication of the experiments conducted (Brygadyrenko, 2017). Please note that at the moment the scientific community is still not so categorical in regards to pseudoreplication of experimental research. Discussions are being conducted on the issue of reality and contrivedness of the problem (Rosenberg, 2019; Kozlov and Hulbert, 2006; Rosenberg and Gelashvili, 2008).

We proceed from the assumption that the biology teaching methods cannot stay on the sidelines of the problems actively discussed in biological science. Moreover, this question lies in the plane of the science methodology. The mastery of methodological knowledge and the ability to apply them is the basis for the formation of a system of biological knowledge for senior high school students. The author's early works were devoted to this question (Komarova, 2017). So, we consider the question of to what extent in school experimentation in biology it is necessary to take into account the requirements that S. Hurlbert identified as the problem of pseudoreplication of experimental research in science, as definitively solvable in the direction of their observance. At the same time, given that the educational subject still differs from the basic science in that it is a didactically adapted version of it, it is necessary to achieve a double effect in organization of a school biological experiment. The first effect is that the results of the educational experiment should be maximum consistent, obtained by true replication. The second effect is that the use of true replication should be maximum ergonomical. Ergonomic in time, cost and complexity.

The purpose of this article is to demonstrate the capabilities of a school model experiment in studying

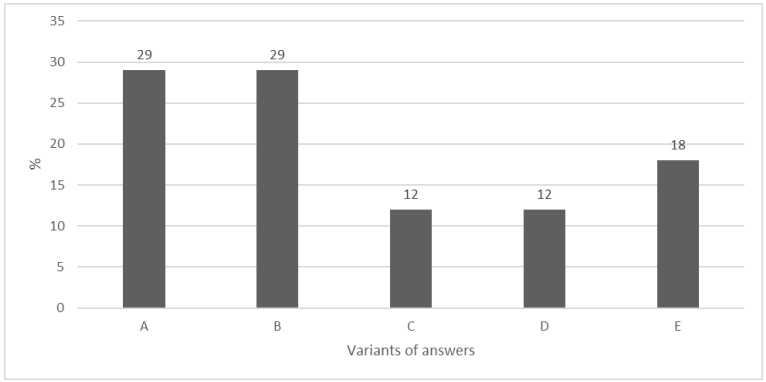


Figure 3: The results of the response to the third task.

the genetic structure of populations over time while meeting the requirements of technical and biological replication. Or, in another way, consider the possibilities of solving the problem of the limitations of an educational model experiment for the study of genetic evolutionary processes in populations.

2 TECHNIQUE AND METHODS

We consider it logical to state the essence of the declared problem in the sequence of answers to the following questions: what is the essence of replication in biological research? What is the essence of pseudoreplication in biological research, what is the history of the problem? How to ensure true replication in a model experiment by studying complex ideal and real biological objects in school biology (by the example of genetic structure of the population)?

Main issue. Why should we conduct at school a model experiment in study of the genetic structure of the population?

Study of the issue of model experimentation with the genetic structure of the population during 2015-2020 convinced us that its goal is the obtention by the students of a direct subject and mediated activity result. Subject result – 1) mastery of the essence of the law of genetic balance and the conditions under which it is consistent; 2) understanding of the mechanism of influence of evolutionary factors, such as natural selection, gene drift, gene flow, mutation process on the genetic structure of the population; understanding of the mechanisms forming the basis of micro- and macroevolutionary processes. The fundamental significance of the law of genetic equilibrium (Hardy-Weinberg) is that it is the central law of population genetics, it is based on the application of statistical methods in genetics (Dodge, 2008).

Before developing a methodology for a model experiment to study the genetic structure of a population, we posed two auxiliary questions:

1. Is it necessary to conduct a model experiment when studying the law of genetic equilibrium and the conditions for its consistency?
2. Can a model experiment be replaced with other educational methods?

The answer to the first question is: no, not necessarily. It is possible to limit to the demonstration of the multimedia presentation and video on this topic.

The answer to the second question is: yes, it is possible. An alternative is familiarization with theoretical material on a printed basis about the factors of change in the genetic structure of a population, overlearning of Hardy-Weinberg equations, teaching of the solution of problems on determination of the genetic structure of a population.

The answers to both questions demonstrate that in the alternative version, at the best, only one result will be achieved – a subject one. Without performing experimental actions, it is extremely difficult to form such elements of methodological knowledge as a variant of experience, replication, sampling. In addition, it has to be considered that the law of genetic equilibrium is a law, the substantial part of which consists of abstract categories not attached to a specific biological object (abstract homozygotes and heterozygotes, dominant and recessive alleles, conditions for the veracity of the law). And the law itself is applicable to some really non-existent ideal object, or, conversely, is not applicable to any really existing object (real population).

The abovementioned reasons are the answer for us to the main question, namely:

1) model experimentation contributes to the mastery by the students of abstract biological categories on concrete material objects;

2) allows to visualize the processes in an ideal population non-existent in reality;
3) allows to simulate the changes taking place in real populations over several generations. Thus, for educational purposes, the time frame of the actually occurring processes is condensed;
4) allows to vary the replications and variants of the experiment with minimal material costs;
5) allows to teach true replicates (replications) of the experimental impact;
6) allows to artificially quickly change the conditions (factors) affecting the population, including acting stochastically (Pavlotsky and Suslin, 1994).

What is the essence of replication in biological research?

A person possessing a basic level of biological knowledge within the scope of the school curriculum of complete secondary education, the term "replication" is known as a process related to the molecular level of organization of a living being. In the English-language scientific biological literature, the term "replication" is used not only in the meaning of the synthesis of new nucleotide sequences, but also in the meaning of the replicate of experimental attempts. In other words, the principle of replication in experiment is the well-known principle of replication. The last term is more widely used in domestic scientific works.

Replication in biological research can be technical and biological (Starmer, 2017).

Technical replicates give us these things:

1. They give us an accurate measurement they give this particular object.
2. If we want to tell more about this object or we do not want to generalize the data and transfer it to the population - a technical experiment is what we need.
3. They will also tell us how accurately we performed the measurements.
4. "If we wanted to publish a paper about how awesome our new method is, we'd use technical replicates" (Starmer, 2017).
5. If the experimental technique is transformed, different samples are taken simultaneously from one object, then technical replication will also take place, since they tell us about an individual.

In the biological replicates each measurement comes from different sample that comes from different objects.

Biological replicates give us these things:

1. Biological replicates tell us about a trait that occurs in a group. In biological replicates, each measurement comes from different samples or is obtained differently from one object.
2. You can mix biological and technical replicates, but the wisdom of doing this depends on the type of the experiment. Sometimes you get more bang for your buck if you add more biological replicates and ignore technical replicates.

   So, the difference between technical and biological replication is as follows: technical replicates are just repetition of the same experiment on the same person.
3. Biological replicates use different biological sources of samples (i.e. different people, different plants, and different cell lines) (Starmer, 2017).

When choosing the type of replication of a biological experiment, it is necessary to proceed from the purpose in view. If it is planned to describe a specific object, whether it be an individual, a population, or to research a method, it is necessary to use technical replication. If the goal is to study a group of objects, it is necessary to choose biological replication.

What is the history of the problem of pseudoreplication in biological research?

The problem of pseudoreplication was raised for the first time in 1984 by S. Hurlbert, who published a critical analysis of 156 experimental scientific papers in English-language editions published in 1960–1980s. He came to the conclusion that in 27% of cases there was one of two variants: 1) the experimental influence was applied in one replication; 2) the experimental replications were not statistically independent. Such errors were called pseudoreplication by S. Hurlbert. M. Kozlov notes that in Russian academic journals in 1998-2001 the part of papers based on pseudoreplication turned out to be twice as high (47%) than in the English-language periodicals for 1960-1980, i.e. before the publication of S. Hurlbert's paper. This situation was considered as non-normal, at the same time it was pointed out that the reason for the pseudoreplication lies not only in errors in experiment planning, but also in the incorrect application of statistical analysis to the results of a well-planned experiment (Kozlov, 2003).

After the publication of S. Hurlbert's paper in 1984 during the period from 1987 to 2001, according to M. Kozlov: 1) the term "pseudoreplication" firmly came into the ecological scientific lexicon of foreign authors, the problem of pseudoreplication in foreign ecological studies is actively discussed; 2) the number of foreign publications based on pseudoreplication began to decrease.

Back in 2003, M. Kozlov paid attention to the fact that the concept of pseudoreplication is completely unknown to the overwhelming majority of Russian ecologists. In addition, the author emphasized that S. Hurlbert's work was never cited in Russian-language periodicals, against the background of more than 2000 references (2015 references as of 2001) in English-language publications. M. Kozlov repeatedly published his works on standing up for the position that the problem of pseudoreplication is a problem of the world scientific community, which should be treated with all possible seriousness (Kozlov and Hulbert, 2006; Kozlov, 2003).

The English term "pseudoreplication" does not have a direct analogue in Russian, since it primarily denotes a process – an erroneous choice of replicates for assessment of intragroup variability in statistical analysis (Hurlbert, 1984; Kozlov and Hulbert, 2006). In this regard, direct translation of terminology is difficult enough; the authors provide English equivalents of key concepts. "In medical experiments, where they are designated to as "spurious replication", "trial inflation", or "the unit of analysis problem or error" (Whiting-O'Keefe et al., 1984; Andersen, 1990; Altman, Bland, 1997). Although the concept of "pseudoreplication", which is most adequately translated as "statistical analysis based on pseudoreplication", is not found in all works listed above, and we do not agree with all the conclusions of the indicated authors, all the cited studies are united by a serious approach to the problem" (Kozlov and Hulbert, 2006).

Is pseudoreplication as scary as it might seem?

In Russian-language sources, the attitude to the problem specified by S. Hurlbert and supported by M. Kozlov can be characterized as far-fetched and already well-known and studied (V. Nalimov, A. Lyubishchev, A. Bakanov, N. Plokhinskiy, T. Golikova). The Russian-speaking authors agree that there are two indisputable theses in the ideas of S. Hurlbert:

1) "it is not always correctly to extend the conclusions, obtained in the study of private samplings, to the entire general population;
2) assessment of the degree of factor influence may turn out to be erroneous if the studied effect is not properly localized, and the compared data are taken from insufficiently randomized sources" (Rosenberg and Gelashvili, 2008).

The conducted analysis of literary sources (Reinhart, 2015; Davies and Gray, 2015; Tatarnikov, 2005; V. Veličković, 2007; Hurlbert, 2004; Oksanen, 2004; Heffner et al., 1996) on the problem allowed us to single out the "pros" and "cons" of the consideration of the problem of pseudoreplication as significant for biological research. The analysis results are presented in table 1.

So, how to ensure replication in a model experiment on population genetics? How to overcome the limitations of the educational model experiment and comply with the conditions for obtaining reliable results?

3 RESULTS

3.1 The Results of the Theoretical Stage of the Study

The development of a model experiment methodology aimed at the study of the essence of genetic-evolutionary changes in the population by students, and its improvement during 2015–2020, was carried out by us in a staged manner. This was dictated by the objective and subjective difficulties of implementation of a model experiment into teaching practice.

At the first stage, we used only material models of gene alleles, created models of genotypes in a manual way, and, respectively, models of parental and daughter populations in generations. Mathematical calculations were performed without the use of a computer, the participants in the experiment manually calculated the frequencies of genotypes and alleles in populations, and presented the results obtained in the graphical representation.

At the second stage, we combined material modelling and use of the computer. Work with material models consisted of carrying out of the experiment itself, creation of a model of the parental population in manual way, and combination of the gene alleles at random (this is how panmixia was simulated). The participants entered the results of the experiments into a table on the developed web pages. With the help of a computer, the obtained frequencies of alleles and genotypes were automatically calculated. In automatic mode, the results of the experiment were optionally presented in the graphical representation.

We have developed a web page for entering, processing and presenting graphical view of modelling results of genetic and evolutionary processes in ideal populations, which are not influenced by the factors of changing its genetic structure (according to the law of Hardy-Weinberg) – http://mybio.education/mod/exp1/en/index.html (Model experiment 1. Study of the genetic structure of the ideal population) and http://mybio.education/mod/exp2/en/index.html (Model experiment 1. Study of the genetic structure of the ideal population (second option)), as well as web pages to make for entering the results of

Table 1: Pseudoreplication – a real problem in biological research.

"Pro" arguments	**"Con" arguments**
1. Each object in the sampling is a functional part of the whole, and not a separate element of a set. In a number of studies, the results and conclusions obtained for discrete objects apply to the entire population, which does not correspond to one of the requirements for biological experimentation – consistency in essence. 2. During the experiment, there is a multiple determination of reaction of the same organism in the course of sampled counts. As an alternative, the same sampling is studied in different time intervals. In this case, living objects (their populations) are pseudoreplications. 3. Two main problems of pseudoreplication is an insufficient mass nature of experimental objects and their initial incomparability with each other. In the first case, the researcher receives insufficient data for the consistent statistical result. In the second case, the problem has an objective causality due to the initial uniqueness of living objects.	1. Each object in the sampling is discrete and individual. 2. Factors acting independently on the sampling, act on a set of separate biological objects, and not on an integral object. The specificity of a biological experiment lies in the uniqueness of the objects and, in certain cases, in the impossibility of repeating the experiment in an accurate manner. 3. Living objects react to the actions of factors independently on a physical level, and thus they are statistically independent. In a majority of research variants, living objects are true replications. 4. The specificity of living objects in their uniqueness and originality. Some ecological research involves study of the reactions of individuals or parts to the impact. In a number of studies, it is not possible to repeat a unique biological object, whether it be an individual or a population. 5. The problem of pseudoreplication is artificial, since technical and biological replication is distinguished in biology. The attempt to apply the goals and requirements of technical replication to biological is a prime cause of the issue of pseudoreplication in biological experiments. 6. According to one of the points of view, the attention of English-speaking authors to the problem of pseudoreplication is explained by several reasons: • the desire to join the campaign of criticism and to incriminate colleagues in pseudoreplication; • the attempt to divert the stigma of pseudoreplication from their work and the work of colleagues; • as a warning signal to the reviewer that the author is acquainted with the work of S. Hurlbert, therefore there should be no comments on the paper (Rosenberg, 2019).

modelling of genetic and evolutionary processes in populations, which are influenced by the factors of changing its genetic structure http://mybio.education/mod/exp3/en/index.html (Model experiment 2. Study of the genetic structure of the population under the influence of natural selection), http://mybio.education/mod/exp4/en/index.html (Model experiment 3. Modelling the effect of gene flow on the genetic structure of the population), http://mybio.education/mod/exp5/en/index.html (Model experiment 4. Modelling the effect of random processes on the genetic structure of the population, modelling the drift of genes).

The system for on-line processing of modeling results developed in 2015 can be used only if the number of model individuals in the population is insignificant in the model experiment. The population size is limited by the objective possibility of creating a corresponding number of chip patterns of the alleles of a gene. Optimum number of chips – 100. In this case, the number of individuals is equal to 50. One can take more or fewer objects. In the first case, the choice will be associated with the growth of material costs for the manufacture of model elements. In the second case, the calculated values (allele frequency) will be significantly deviate from the pre-selected frequencies, and the level of statistical significance of the obtained results will decrease.

Stages of modelling of the genetic structure of populations are as following:

1. Modelling of the genetic structure of an ideal population with the use of material objects. Entry of simulation results into a table on web pages http://mybio.education/mod/exp1/en/index.html or http://mybio.education/mod/exp2/en/index.html.

Modelling of the genetic structure of an ideal population can be done using the possibilities of any of the two web pages. The difference between them lies in the methods of processing of the experimental results, namely in the methods of calculating the frequencies of genes. In the first variant, the gene frequencies are calculated automatically by the method of extracting of the square roots of the frequencies of the homozygotes AA and AA. In the second variant the gene frequencies are automatically calculated according to the formulas: $p = (D + 0.5H)/N$, $q = (R + 0.5H)/N$, where p – frequency of dominant allele, q – frequency of recessive allele, D – number of dominant homozygotes, R – number of recessive homozygotes, H – number of heterozygotes, N – total number of members of the population. Both methods allow us to formulate the main conclusion, that in ideal populations, the ratio of frequencies of genes and genotypes remain constant from generation to generation, and the sum of their frequencies is equal to 1.

2. Modelling of population genetic structure, which is influenced by factors of change in its genetic structure – natural selection, gene flow, genetic drift. Entry of simulation results into a table on web pages http://mybio.education/mod/exp3/en/index.html, http://mybio.education/mod/exp4/en/index.html, http://mybio.education/mod/exp5/en/index.htmlrespectively.

Before usage of web pages for entering the results of the simulation, high school students work with persisted models of alleles of dealing a gene and create a genetic model of the parent population. These materialized models can be checkers, chips, candies, balls of different colours. The educational models of the genetic structure of the population are the findings of the experimental action with the model elements first ratio of genotypes and ratio of frequencies of genes, that is, the ratio of frequencies of genotypes and genes in the parent population.

On each of the web pages there is an instruction for the sequence of actions that must be performed concerning materialized objects, as well as actions to enter the received results in the tables for automatic calculation of genotype frequencies and allele frequencies. The rows that are highlighted in blue in tables for web pages http://mybio.education/mod/exp1/en/index.html, http://mybio.education/mod/exp2/en/index.html, http://mybio.education/mod/exp3/en/index.html or http://mybio.education/mod/exp4/en/index.html, http://mybio.education/mod/exp5/en/index.html are filled manually by students on the basis of counting of the number of the results obtained in the course of the materialized models of alleles and genotypes. The web pages provide automatic plotting of graphs and charts, allowing, first, to reveal the results in graphical form (figures 4, 5). Secondly, it allows to effectively carry out their comparative analysis and to formulate conclusions according to the algorithm of the action plan.

Both diagrams show the genetic structure of populations and according to the semantic content they are identical. They differ in the way of the visibility of the results. The teacher can draw the students' attention to one variant of a diagram with a proposal to compare the genetic, genotypic structure of the population in generations. There is another, more complicated version of the analysis of the constructed diagrams. For this the students choose their own chart to analyze data and formulate conclusions.

Both variants have advantages and disadvantages. In the first variant of the diagram, numeric data of the results of the experiment are included in the corresponding segments of each column. All the data are displayed on the screen, so the student can quite easily compare the numbers.

In the second variant, the segments of each column are located one behind the other, and so that the first, the most narrow segment corresponds to the parent generation and the last, the widest one corresponds to the last child generation. This way of presenting data is liked by students because, not even using numerical data it is visually easy to compare the size (height) of colored bars. Besides, when one aims the cursor at the corresponding field the necessary numerical information appears on the screen.

Analysis of the received data of the model experimentation by the students is carried out on the basis of the analysis of the built:

1) graphics of genetic and genotype structure of the population in generations;
2) one of the diagrams of the genetic structure of the population in generations;
3) graphs and diagrams that overlap.

A variety of graphic options allows to acquaint students with the methods of their statistical processing and presentation.

Testing of the developed web pages and work on the proposed methods during 2015–2018 demonstrated that the proposed options did not allow working with a large number of experimental objects. That is, it was impossible to comply with the condition of the mass scale of the experiment. In addition, the question of the replicativity of the experiment also remained open. The reasons are as follows:

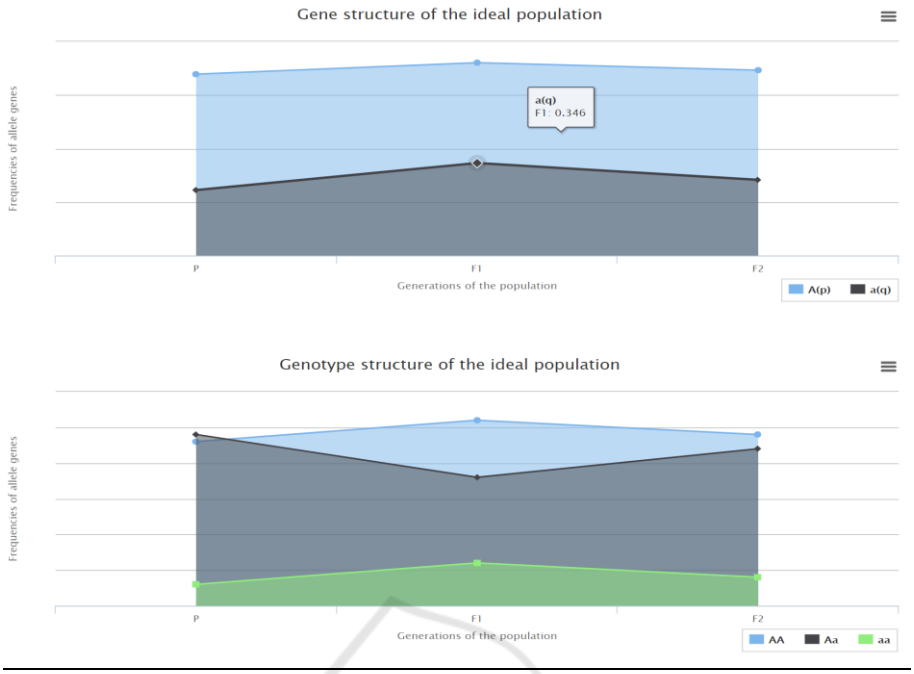


Figure 4: View of graphics on the web page that were built in automatic mode http://mybio.education/mod/exp1/en/index.html.

1. It is physically impossible in the course of the educational process to explore a large number of material model objects – homozygous dominant, homozygous recessive and heterozygous individuals. The work was accompanied by the enormous time spent on manual modelling and counting of randomly formed pairs of alleles. Such a calculation had to be carried out both within one generation, and in several replications. Note that in this variant we are talking about the difficulties with the technical replication of the experiment.
2. The use of material models was limited to elementary material costs for the manufacture of model elements. The maximum number of individuals whose genotype models were used in the experiment was equal to 50. In the case of diallelic inheritance of a trait (as the simplest variant of inheritance), the number of alleles was equal to 100. Let's point to the fact that in the classrooms there were carried out parallel experiments on the study of influence of different factors of the dynamics of the population genetic structure, the work was carried out in small groups, each of which worked with a separate set of elements for modelling. There were 5 such groups. The first group studied the genetic structure of an ideal population in generations. The second group studied the effects of gene drift. The third group studied the essence of the gene flow phenomenon. The fourth group studied the influence of natural selection on the genetic structure of the population. The fifth group studied the role of the mutational process in the dynamics of the genetic structure of the population. In total, at least a set of 500 material elements was needed for modelling.

We place the emphasis on the fact that even with 50 simulated members of the population, we obtained results that allowed to illustrate the essence of genetic transformations in populations in the absence of any factors and in their presence.

In work on the improvement of the experimental methodology, we tried to: 1) get closer in school modelling of genetic-evolutionary processes to the real process taking a course in populations; 2) take into account significant differences and commonality between scientific and educational experiment. Particularly, this was expressed in the fact that it was necessary to:

1. Cover by the experiment the maximally large number of individuals. It has been assumed that the hundreds and thousands of individuals could be the experimental objects.
2. Reduce the amount of routine work for students on the calculation of the resulting genotypes and

Figure 5: View of diagrams on the web page that were built in automatic mode.

alleles in one generation.

3. Simulate a larger number of replications (replicates) of the experiment, which would increase the veracity of results and their closeness to the mathematical formula of Hardy-Weinberg. We also set the task to provide the possibility for technical and biological replication of an experiment on one topic.

Thus, we set the task to provide the possibility of technical and biological repetition of an experiment on one topic.

Taking into account the abovementioned tasks, we have developed a web page http://mybio.education/mod/exp6/en/index.html (Model experiment 1. Study of the genetic structure of the ideal population (third variant).

Using the tools of this web page, we can conduct an experiment on the modelling of a structure of an ideal population in the absence of such factors as natural selection, gene flow, gene drift, mutations. Note that this is the third variant for conduction of a model experiment on the stated topic. The first two are displayed on the following pages: http://mybio.education/mod/exp1/en/index.html and http://mybio.education/mod/exp2/en/index.html.

What is the difference between the proposed third variant?

First. In two early variants, the number of individuals was limited by the physical ability to manually count the resulting pairs of alleles and the number of material elements for modelling. The studied population in the proposed variant can be very large – several hundreds, thousands, millions of individuals. This contributes to the implementation of the first of the tasks pursued by us – an increase in the number of objects used in the model experiment. And in this case, it can be considered as a step towards the increase in the veracity of the experimental results. And thus, the maximum convergence with the actually occurring genetic and evolutionary transformations in the population. For example, for an experiment, you can take several tens of thousands of individuals and several million (figures 6, 7).

Second. The user (student) can independently enter the initial allele frequencies in the graph for the parental population. In the first two variants, the allele frequencies were calculated automatically after data entering by the manual calculation of the ran-

Model experiment 1. Study of the genetic structure of the ideal population (third option)

1. In column 2 for the parent generation P, we introduce the number of pairs of two-body gene alleles (in other words, the number of individuals).
2. We are determined by the ratio of the dominant (A) and recessive (a) alleles, and we introduce their values in columns 9 and 10 for the parent generation P.
3. Click the "Calculate" button.
4. Click the "Calculate" button opposite the lines F1, F2, F3, F4, F5.
5. Click on the "Show Graphs" button.
6. Based on the analysis of the obtained graphs and diagrams, formulate the conclusions of the plan:
   - Change in the frequency of genotypes in generations;
   - Change in the ratio of gene frequencies in generations;
   - The direction of evolutionary changes in the population.

Table 6. Genetic structure of the ideal population

Generation	Number of individuals	Distribution of genotypes						Gene frequencies		
		AA		Aa		aa		A(p)	a(q)	
1	2	3	4	5	6	7	8	9	10	
P	20000	10910	0.545	6180	0.309	2910	0.145	0.7	0.3	Calculate
F1	20000	10893	0.545	6214	0.311	2893	0.145	0.738	0.380	Calculate
F2	20000	10883	0.544	6234	0.312	2883	0.144	0.738	0.380	Calculate
F3	20000	10890	0.544	6220	0.311	2890	0.144	0.738	0.380	Calculate
F4	20000	10823	0.541	6354	0.318	2823	0.141	0.736	0.376	Calculate
F5	20000	10869	0.543	6262	0.313	2869	0.143	0.737	0.379	Calculate

Show Graphs

Figure 6: Results of a model experiment with a number of 20000 individuals, allele frequencies p (0.7) and q (0.3).

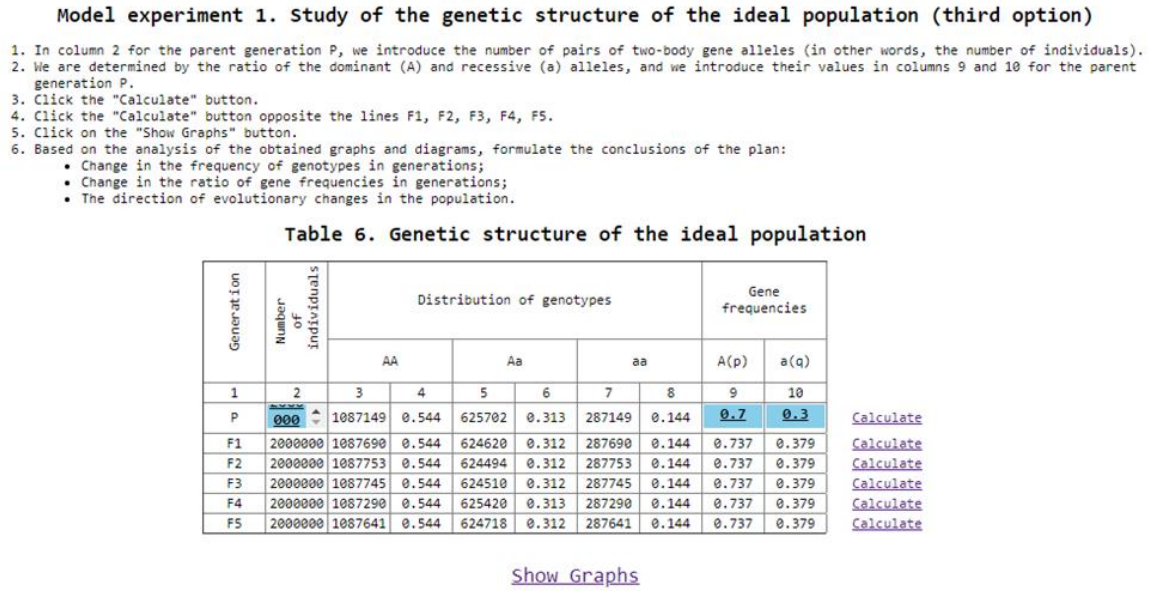
domly obtained genotypes. This function opens up an opportunity to demonstrate the essence of **biological replication of experiments**. This function is especially remarkable in the lesson during the simultaneous work of several groups of students with different populations in number and frequency of alleles occurrence (figures 6, 8).

Third. The number of generations of the population has been increased. In the proposed variant, it is equal to 5. I.e. together with the parental population, the total number of replications of the experiment is equal to 6. In previous variants of the experiment, the number of replications was equal to 3 (one parental generation and two daughter generations). In addition, note that it is technically possible to increase the number of replications by times. This will offer an opportunity, first of all, to quickly get a picture of the genetic structure of the population, without bothering students with mechanical work on mixing and distribution of genotypes, since there is an automatic distribution of genotype frequencies within the limits of the ideal population. Secondly, it contributes to the implementation of one of the tasks pursued by us – an increase in the number of replications of the experiment within the limit of one sample (population). This function opens an opportunity to conduct the **technical replication of experiments**.

The visualized replication results are displayed on the user's screen by clicking the "Show graphs and diagrams" button. Note that in one session the user can only see the results of technical replication, i.e. distribution of alleles and genotypes in generations with initially specified parameters (number of individuals and allele frequencies). The generations of the population will act as technical replications. In order to simulate biological replication, it is necessary to load the page once again without closing the previous one and enter other initial data (the number of individuals, allele frequencies). Within each session, generations of a population in relation to each other will act as technical replications, but in relation to the first population and its generations – biological replications.

3.2 The Results of the Experimental Stage of the Study

In 2019/2020 academic year, the developed web page was tested with the participation of 6 students of the 3rd year of the Institute of Living Systems of Immanuel Kant Baltic Federal University, in specialty "Biology" and 12 students of the 11th form of the Municipal Budgetary General Education Institution General Secondary School "School of the Future" of Guryevsky district of Kaliningrad Region (Russian Federation). The approbation took place within the framework of carrying out by Municipal Budgetary General Education Institution General Secondary School "School of the Future" together with the National Research University Higher School of Economics of the conference "Effective High School" (January 23–25, 2020). Within the framework of the conference, there were organized practical classes for



Model experiment 1. Study of the genetic structure of the ideal population (third option)

1. In column 2 for the parent generation P, we introduce the number of pairs of two-body gene alleles (in other words, the number of individuals).
2. We are determined by the ratio of the dominant (A) and recessive (a) alleles, and we introduce their values in columns 9 and 10 for the parent generation P.
3. Click the "Calculate" button.
4. Click the "Calculate" button opposite the lines F1, F2, F3, F4, F5.
5. Click on the "Show Graphs" button.
6. Based on the analysis of the obtained graphs and diagrams, formulate the conclusions of the plan:
   - Change in the frequency of genotypes in generations;
   - Change in the ratio of gene frequencies in generations;
   - The direction of evolutionary changes in the population.

Table 6. Genetic structure of the ideal population

Generation	Number of individuals	Distribution of genotypes						Gene frequencies		
		AA		Aa		aa		A(p)	a(q)	
1	2	3	4	5	6	7	8	9	10	
P	000	1087149	0.544	625702	0.313	287149	0.144	0.7	0.3	Calculate
F1	2000000	1087690	0.544	624620	0.312	287690	0.144	0.737	0.379	Calculate
F2	2000000	1087753	0.544	624494	0.312	287753	0.144	0.737	0.379	Calculate
F3	2000000	1087745	0.544	624510	0.312	287745	0.144	0.737	0.379	Calculate
F4	2000000	1087290	0.544	625420	0.313	287290	0.144	0.737	0.379	Calculate
F5	2000000	1087641	0.544	624718	0.312	287641	0.144	0.737	0.379	Calculate

Show Graphs

Figure 7: Results of a model experiment with a number of 2000000 individuals, allele frequencies p (0.7) and q (0.3).

students of 11th grade on the topic "Modelling of the genetic evolutionary processes in the population". One of the proposed experiments for carrying out was a model experiment "Study of the genetic structure of an ideal population" according to the methodology updated by us without using material objects.

In approbation, the participants were divided into 2 groups (3 students and 6 students). One group was asked to start with an experiment at http://mybio.education/mod/exp1/en/index.html, and then at http://mybio.education/mod/exp6/en/index.html. Another group was asked to click a link to the web page http://mybio.education/mod/exp6/en/index.html (Model experiment 1. Study of the genetic structure of the ideal population (third variant) and simulate the genetic structure of a population of any number more than a thousand with an arbitrarily given combination of allele frequencies. It was proposed three times to the participants to carry out model experiment in the third variant with different initial data (number of individuals of the population, allele frequencies). Each of the participants of the approbation both in the first and second groups in carrying out of the third variant of the experiment worked separately. The participants were asked to use the "Show graphs and diagrams" function, and also to formulate conclusions at the end of the experiment.

The goals pursued by us were as follows:

1. To find out the availability of understanding by users of the tasks and results of experiments.
2. To find out the main difficulties faced by users when working with a web page http://mybio.education/mod/exp6/en/index.html.

During the oral survey of the participants in the experiment, it was found:

1. Participants of the first group, when conducting an experiment with material objects at the beginning of work, hardly understood the essence of the performed similar actions. Only after data entering into the table, calculation of the frequencies of alleles and genotypes, the understanding of the meaning of the uniformity of actions came.
2. Participants of the first group complained about the routine of the performed actions, increased fatigue during their performance. Participants sought to complete the experiment more quickly, which increased the error rate in calculation of the absolute number of genotypes. The latter was displayed at the frequency of genotypes calculated by the program. Thus, the obtained results in several cases were erroneous, the experimental actions had to be performed over again.
3. Participants of the first group, after passage to the second experiment, which, in fact, duplicated the first variant, but did not require manual counting, expressed great approval of the possibility to operate only with numbers.
4. Participants of the second group completed the assigned task more quickly. However, in both groups, there arose questions about the purpose of three-time replicate of the experiment (with dif-

Model experiment 1. Study of the genetic structure of the ideal population (third option)

1. In column 2 for the parent generation P, we introduce the number of pairs of two-body gene alleles (in other words, the number of individuals).
2. We are determined by the ratio of the dominant (A) and recessive (a) alleles, and we introduce their values in columns 9 and 10 for the parent generation P.
3. Click the "Calculate" button.
4. Click the "Calculate" button opposite the lines F1, F2, F3, F4, F5.
5. Click on the "Show Graphs" button.
6. Based on the analysis of the obtained graphs and diagrams, formulate the conclusions of the plan:
   - Change in the frequency of genotypes in generations;
   - Change in the ratio of gene frequencies in generations;
   - The direction of evolutionary changes in the population.

Table 6. Genetic structure of the ideal population

Generation	Number of individuals	Distribution of genotypes						Gene frequencies		
		AA		Aa		aa		A(p)	a(q)	
1	2	3	4	5	6	7	8	9	10	
P	20000	2295	0.115	3410	0.171	14295	0.715	0.2	0.8	Calculate
F1	20000	2261	0.113	3478	0.174	14261	0.713	0.336	0.844	Calculate
F2	20000	2310	0.116	3380	0.169	14310	0.716	0.340	0.846	Calculate
F3	20000	2274	0.114	3452	0.173	14274	0.714	0.337	0.845	Calculate
F4	20000	2221	0.111	3558	0.178	14221	0.711	0.333	0.843	Calculate
F5	20000	2285	0.114	3430	0.172	14285	0.714	0.338	0.845	Calculate

Show Graphs

Figure 8: Results of a model experiment with a number of 200000 individuals, allele frequencies p (0.2) and q (0.8).

ferent number of population and allele frequencies). Note that practically no questions arose in both groups regarding the advisability of repeating the experiment in generations of the same population. It follows that the essence and necessity of technical replication is recognized and accepted by the participants.

With biological replication, the situation is different. Its objectives were not clear to the participants, most likely due to a lack of methodological awareness of this type of replication.

Before the performance of the experiment, we deliberately did not focus the participants' attention on the goals of repeated replicate of experimental actions. This was done in order to find out whether the participants understood the conditions for the veracity of the results of the biological experiment. Since among the examinees there were both students of a biological specialty and students of graduating profile chemical and biological classes. We assumed that the participants already possess the necessary methodological tools for planning, conduction of biological experiments and interpretation of the results. The results of approbation showed that teaching the methodology of a biological experiment should be started with distinguishing between technical and biological replication of experimental effects. We can only assume that a lack of understanding of the differences between them (for purposes, methodology) could initiate the spread of the problem of pseudoreplication in biological research in principle. We believe that in order to confirm this assumption, it is necessary to conduct additional studies aimed at a retrospective analysis of biological scientific literature, primarily of scientific papers, conference materials containing a description of the methods and results of experiments. The question is, is it worth doing? Or to accept the fact that even if we consider the problem of pseudoreplication as far-fetched, then the issue of distinguishing between technical and biological replications and teaching this in the secondary school and in higher educational establishment definitely deserves further study.

4 CONCLUSIONS

As a result of work on the topic of overcoming methodological difficulties in conducting a model experiment on population genetics, we came to the following conclusions:

1. Model experiment on the study of genetic-evolutionary processes in populations by means of computer modelling is ideal for demonstration of the essence of technical and biological replication.
2. The use of technical and biological replication in a model experiment makes it possible to take into account the requirement of repetition and mass scale of the experimental impact. This is neces-

sary to obtain reliable results.

3. In the educational model experiment, it is impossible to take into account all the requirements for a scientific biological experiment, therefore, it is necessary to rely only on its essential features.

5 OUTLOOK

Further work on studying the possibilities of a model experiment in training of the students of 11 grade and students-biologists in true replication, as well as the essence of technical and biological replication, we can see in the following. It is necessary to develop and approbate web pages to model the structure of a very large population under the influence on its numerous generations of such factors as natural selection, gene flow, gene drift, mutations.

The modelling of the genetic structure should be fully automated. The initial platforms for improvement of methodology will be the existing web pages http://mybio.education/mod/exp3/en/index.html (Model experiment 2. Study of the genetic structure of the population under the influence of natural selection), http://mybio.education/mod/exp4/en/index.html (Model experiment 3. Modelling the effect of gene flow on the genetic structure of the population), http://mybio.education/mod/exp5/en/index.html (Model experiment 4. Modelling the effect of random processes on the genetic structure of the population, modelling the drift of genes*), providing one of the stages of work with material objects.

REFERENCES

Bazykin, A. D. (1988). Modelirovanie biologicheskikh protcessov (Modeling of biological processes). *Biology at school*, (4):5–9.

Binas, A. V., Mash, R. D., and Nikishov, A. I. (1990). *Biologicheskij eksperiment v shkole (Biological experiment at school)*. Prosveshchenie, Moscow.

Borodin, P. M. (1987). Modelnye eksperimenty po genetike i evoliutcii populiatcii (Model experiments on genetics and population evolution). *Biology at school*, (1):49–53.

Brunovt, E. P., Zverev, I. D., and Malakhova, G. I. (1973). *Metodika prepodavaniya anatomii i fiziologii cheloveka (Methods of teaching human anatomy and physiology)*. Prosveshchenie, Moscow.

Brygadyrenko, V. V. (2017). Kak opublikovat statiu v Scopus? (How to publish an article in Scopus?). https://openscience.in.ua/scopus-article.html.

Bulaeva, K. B. (1977). Izuchenie zakona Khardi-Vainberga v kurse obshchei biologii (Study of Hardy-Weinberg's law in the course of general biology). *Biology at school*, (6):46–49.

Davies, G. M. and Gray, A. (2015). Don't let spurious accusations of pseudoreplication limit our ability to learn from natural experiments (and other messy kinds of ecological monitoring). *Ecology and Evolution*, 5(22):5295–5304. https://onlinelibrary.wiley.com/doi/abs/10.1002/ece3.1782.

Dodge, Y. (2008). *The Concise Encyclopedia of Statistics*. Springer, New York.

Frolov, I. T. (2007). *Ocherki metodologii biologicheskogo issledovaniya: Sistema metodov biologii (Essays on the methodology of biological research: system of methods of biology)*. LKI, Moscow.

Hammersmith, R. L. and Mertens, T. R. (1990). Teaching the concept of genetic drift using a simulation. *The American Biology Teacher*, 52(8):497–499. http://www.jstor.org/stable/4449186.

Heffner, R. A., Butler, M. J., and Reilly, C. K. (1996). Pseudoreplication revisited. *Ecology*, 77(8):2558–2562. https://esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/2265754.

Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, 54(2):187–211. https://esajournals.onlinelibrary.wiley.com/doi/abs/10.2307/1942661.

Hurlbert, S. H. (2004). On misinterpretations of pseudoreplication and related matters: a reply to Oksanen. *Oikos*, 104(3):591–597. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0030-1299.2004.12752.x.

Komarova, E. and Starova, T. (2020). Majority values of school biological education in the context of education for sustainable development. *E3S Web of Conferences*, 166:10029.

Komarova, O. and Azaryan, A. (2018). Computer simulation of biological processes at the high school. *CEUR Workshop Proceedings*, 2257:24–32.

Komarova, O. V. (2017). *Theoretical and methodical bases of formation of system of knowledge of students in the learning process of general biology*. D.Sc. Thesis, Institute of Pedagogy of the NAES of Ukraine. https://tinyurl.com/h9vnx50u.

Kozlov, M. V. (2003). Pseudoreplication in ecological research: the problem overlooked by Russian scientists. *Zhurnal obshchei biologii*, (64):292–307.

Kozlov, M. V. and Hulbert, S. H. (2006). Pseudoreplication, chatter, and the international nature of science: a response to D.V. Tatarnikov. *Zhurnal obshchei biologii*, (67):145–152.

Maret, T. J. and Rissing, S. W. (1998). Exploring genetic drift & natural selection through a simulation activity. *The American Biology Teacher*, 60(9):681–683. http://www.jstor.org/stable/4450580.

Mertens, T. R. (1992). Introducing Students to Population Genetics & the Hardy-Weinberg Principle. *The American Biology Teacher*, 54(2):103–107. http://www.jstor.org/stable/4449417.

Moore, R., editor (1994). *Biology Labs That Work: The Best of How-To-Do-Its*. National As-

sociation of Biology Teachers, Reston, VA. https://files.eric.ed.gov/fulltext/ED411149.pdf.

Mukhopadhyay, C., Henze, R., and Moses, Y. T. (2014). *Genetic Drift and Gene Flow Illustration. How Real is Race? A Sourcebook on Race, Culture, and Biology*. AltaMira Press, Lanham, 2nd edition. http://www.sjsu.edu/people/carol.mukhopadhyay/race/Gene-Flow-and-Genetic-Drift-Illustration-2014.pdf.

Nechypurenko, P. P., Selivanova, T. V., and Fedorynova, N. Y. (2021). Analysis of some aspects of the implementation of the integrated course "Science" in the educational process of schools in Ukraine. *Journal of Physics: Conference Series*, 1840(1):012037. https://doi.org/10.1088/1742-6596/1840/1/012037.

Oksanen, L. (2004). The devil lies in details: reply to Stuart Hurlbert. *Oikos*, 104(3):598–605. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0030-1299.2004.13266.x.

Pavlotsky, I. P. and Suslin, V. M. (1994). Stochastic model of evolution on areas in population ecology. *Matem. Mod.*, 6(3):9–24. http://mi.mathnet.ru/mm1844.

Pongsophon, P., Roadrangka, V., and Campbell, A. (2007). Counting Buttons: demonstrating the Hardy-Weinberg principle. *Science in School*, (6):30–35.

Reinhart, A. (2015). *Statistics Done Wrong: The Woefully Complete Guide*. No Starch Press, San Francisco. https://www.statisticsdonewrong.com/.

Rosenberg, G. S. (2019). Several comments of the translator about "truth, lies and statistics". *Samarskaya Luka: problems of regional and global ecology*, (28):53–38.

Rosenberg, G. S. and Gelashvili, D. B., editors (2008). *Problemy ekologicheskogo eksperimenta (Planirovanie i analiz nabliudenii) (Problems of ecological experiment (Planning and analysis of observations))*. Cassandra, Togliatti.

Shamrai, S. M. and Zadorozhnyi, K. M. (2003). *Biologichni eksperimenti v shkoli (Biological experiments in school)*. Osnova, Kharkiv.

Sidorova, N. A. (2009). Matematicheskoe modelirovanie pri izuchenii temy "Genetika i evoliutciia populiatcii" (Mathematical modeling in the study of the topic "Genetics and evolution of populations). *Biology at school*, (6):27–29.

Starmer, J. (2017). StatQuest: Technical and Biological replicates, clearly explained! https://statquest.org/statquest-technical-and-biological-replicates-clearly-explained/.

Tatarnikov, D. V. (2005). On methodological aspects of ecological experiments (comments on M. V. Kozlov publication). *Zhurnal obshchei biologii*, (66):90–93.

V. Veličković, M. (2007). Sampling designs, pseudoreplication and a good practice in modern science: a response to Mikhail V. Kozlov desultoriness, and recommendations to environmental scientists. *Hereditas*, 144(1):45–47. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2007.0018-0661.01995.x.

Voronin, L. G. and Mash, R. D. (1983). *Metodika provedeniya opytov i nablyudenij po anatomii, fiziologii i gigiene cheloveka (Methodology for conducting experiments and observations on human anatomy, physiology and hygiene)*. Prosveshchenie, Moscow.