Article

Online Blended Learning in Small Private Online Course

Yong Han [1], **Wenjun Wu** [1,*], **Lijun Zhang** [2] and **Yu Liang** [1]

1   State Key Laboratory of Software Development Environment, School of Computer Science and Engineering, Beihang University, Beijing 100191, China; hanyong@buaa.edu.cn (Y.H.); yliang@buaa.edu.cn (Y.L.)
2   School of Computer Science and Engineering, Beihang University, Beijing 100191, China; ljzhang@buaa.edu.cn
*   Correspondence: wwj09315@buaa.edu.cn

Abstract: In this work, we studied the online blended learning model of computer network experimentation, focusing mainly on the problem of traditional network experiments being limited by location and time, and explore the applicability of the small private online course (SPOC) advanced teaching concepts to computer network online experiment teaching. Based on the structure of a combination of virtual and real, real and not virtual, an online network experiment platform and management system has been designed and constructed, enabling students to carry out remote online computer network hardware experiments anytime and anywhere, without being restricted by time, space, or content. Using the online network experiment platform, we can organize the experimental modules and knowledge points via the SPOC course concept, by developing online network experimental content, modularizing and fragmenting of the experiments, creating the pre-experimental explanation and experiment preview videos, and evaluating the assignments via peer grading to analyze students' learning behavior. By exploring online network experimental teaching methods and management models, offering experimental guidance in an interactive manner, and highlighting the openness and sharing characteristics of online experimental teaching platforms, we can improve the utilization rate for teaching resources, and provide ideas for applied scientific research methods.

Keywords: peer grading; blended learning; online experiment platform

1. Introduction

In this paper we start with the online blended learning method, and study the beneficial effects of online experimental platform construction, small private online course (SPOC) construction, and assignment evaluation on experiment teaching.

The traditional structure of experimental teaching resources has increasingly restricted students' practical learning, and there is an urgent need to resolve these limitations through revolutionary technical means, so that the quality of experimental teaching can be improved. Fortunately, in recent years, due to the development of the mobile internet, cloud computing, big data, and artificial intelligence, a solid technical foundation has been laid for revolutionary breakthroughs in the experimental methods of computer-oriented courses. It is against this background that the traditional computer experimental methods have been reformed using online experimental platforms.

An online experiment platform based on real equipment, as opposed to a simulation using pure software, has obvious advantages. Compared with commercial-grade network equipment, the engineering practice, functional integrity, and reliability of simulation software are obviously insufficient. For example, the experimental results of simulation software cannot be intuitively observed, and it also falters in cultivating students' hands-on ability. In addition, the network protocols of the simulation software (Zebra, Enterprise Network Simulation Platform, H3C HCL) have incomplete and irregular implementations [1]. Therefore, we believe that it is necessary to build a remote online virtual simulation experiment environment composed of real experimental equipment and software platform.

In order to facilitate practical learning for anyone, at any time, and in any place, it breaks through the technical bottleneck of the networked use of experimental hardware, constructs a comprehensive online hardware and software experimental environment and system, and realizes the transformation from a traditional experimental teaching center to an online experimental teaching and research center.

Our network experiment courses use a blended learning method that combines offline experiments with online experiments and online peer grading to improve students' hands-on ability and their level of knowledge. The blended learning method based on SPOC is designed as follows: (1) Online preview. Before each class, students learn about the experimental content, experimental procedures, and experimental precautions by watching our explanation videos. (2) Offline or online experiments. Students can complete the experiment by attending the physics laboratory, or they can make an appointment on the online experiment platform to complete the experiment. (3) Peer grading. After completing the experiment, students need to submit their own process and results online, thus deepening their understanding of experimental knowledge through peer grading between peers.

Finally, we verify that the blended learning method proposed in this paper is indeed effective through an analysis of the learning effect. Through analyses of the teaching method, the submissions (and thus the learning) of the students, and also the students' exam grades, we analyze the gaps in students' knowledge and provide learning path planning. This online blended learning method provides a new way of studying educational science.

In this paper we convert the traditional network experiment environment into an online experiment environment, and combine online experiments and online courses, so that not only can students combine theoretical learning and experiments, but teachers can also collect students' learning trajectory data in order to improve the learning effect. The rest of the paper is organized as follows: Section 2 discusses work related to our research. Section 3 elaborates the construction of the computer network online experiment platform. Section 4 describes the construction of the SPOC course for network experiments. Section 5 describes the analysis of student behavior in the context of an SPOC course on network experimentation. Section 6 presents the experimental results of the student behavior analysis and a verification of the effects of the blended learning methods. Section 7 provides our conclusions.

2. Related Work

Online blended learning integrates an experimental platform into online courses, so that students can use the experimental platform to conduct experiments anytime and anywhere remotely and learn online at any time, combining hands-on operation and theoretical learning.

In the traditional network experiment model, students must go to the laboratory to use the experimental resources. Limited by this mode of use, traditional experimental teaching resources occupy much physical space. Furthermore, due to the restriction of management methods, students can only use the experimental resources at a specified place and time [2,3]. The traditional model struggles to provide fine-grained procedural data for analyses of students' learning behavior, making it difficult to evaluate and quantify students' learning effects in detail.

Many scholars have thought of using pure software simulation for student experiments. Students can only see some images and configurations on the computer, and do not experience the actual physical experiment environment [4]. The network packets intercepted by the students in the network experiment simulation environment are not sent by real network equipment, and differ from data from a real network environment, so the observation results are not intuitive.

Although the use of simulation software has the advantages of low cost and convenient usability, it has shortcomings in cultivating students' practical ability, as many experiments can only be debugged on experimental equipment, rather than in simulation experiments.

Students can successfully complete experiments on a network simulator, but in actual work situations, they often struggle. Often, students know the experimental principles, but are slow to get started, and do not know how to troubleshoot when encountering a fault. The main reason is that the network simulator cannot simulate common faults such as poor contact and unsuccessful agreement negotiation between devices, which leads to a lack of practical and engineering practice ability.

With the development of networks, online learning plays an increasingly important role in promoting education. Online learning has the advantages of convenience, freedom, and low costs [5]. People can learn whatever they want on the internet by just moving their fingers [3]. Physics online experiment platforms are set up on the internet in the same way as an SPOC, meaning students are not restricted by physical conditions such as venue and time, which promotes the continuous growth of experimental teaching resources [6–8]. Therefore, we designed an online experiment platform and an online network experiment SPOC course to improve the effects of network experiment teaching.

The blended learning method not only provides convenience for students to conduct network experiments, but also provides a data basis for analyzing students' learning trajectories. The blended learning method facilitates the further analysis of students' learning behavior, via monitoring the students' interactive behaviors when watching videos and forming teaching research methods based on micro-quantitative analysis. For example, in order to improve the learning effect, some scholars measure the ability of students according to the correctness of students' answers to questions [9,10]. We also set up open questions to encourage students to actively participate in the peer grading of assignments. In the analysis of the reliability of the peer-grader, models such as Piech C [11] and Mi F [12] have been proposed for use in SPOC courses, with the biases and reliabilities of the graders as the major factors. In the context of peer grading tasks, Yan Y [13] proposed an online adaptive task assignment algorithm, thus ensuring the reliability of the results. The algorithm determines redundancy and functionality according to the requirements of the task by analyzing the time taken to complete the task and the confidence in the results, in order to make the final results highly reliable. In their research, Gao A [14] proposed a combination of game-based peer prediction and teacher-provided Ground Truth scores in the peer assessment of assignments.

We propose a blended learning method, which can not only provide students with hands-on ability, but also provide a more convenient and flexible way for students to conduct experiments. It can also offer personalized learning path planning through the analysis of students' learning behavior. At the end of the paper, with reference to specific evaluation methods [15], we assess the effect of the blended learning method proposed in this paper.

3. Construction of Computer Network Experiment Platform

The computer network experiment platform includes a physical experiment environment and an online experiment environment. The physical experiment environment allows students to perform experiments in the laboratory within specified classroom time. The environmental elements and basic configuration of the online experiment environment are exactly the same as those of the physical experiment environment, which addresses the problem of the large number of experiment appointments made during the course time and the unresolved issues students have after the class time is over.

3.1. Physical Experiment Environment for Network Experiment

We set up a number of computers as terminals for the student experiments. We divided these terminals into groups—each group contained four desktop computers, and each was connected through a router and a switch. We configured an ethernet interface and a serial port for each virtual machine, and the serial port was directly connected to the router through the console line.

In order to make full use of wired resources, students had to make an appointment before each experiment. They had to complete their own experiment content within the specified time, and take screenshots to document the process and results. We also gave a grade for the student's attendance and performance during the experiment, which would be included in the student's final grade, in order to encourage the students to complete the experiments properly. During the experiment time, teachers or teaching assistants (TAs) were available to answer questions on the spot.

3.2. Online Experiment Environment for Network Experiment

The computer network online experiment platform is a combination of virtual and real, and real and not virtual, and aims to deploy a real network laboratory on the "Internet". We combined equipment hardware, network communication, multimedia, databases, and other technologies, and through the integration of experimental hardware and planning with an uninterrupted online network environment, students could practice it in their dormitory or at home as if they were in a network laboratory. Through the remote experimental environment, we could also carry out multi-person joint experiments in different places with the remote guidance and assessment of teachers [16].

The overall structure of the online network experimental platform is shown in Figure 1. The entire platform consists of three parts: (1) The online network management system is the core of the entire platform, and is composed of user and authority management, experiment resource management, and experiment reservation modules. (2) The server cluster is composed of several servers and the computing and management resources required for network experiments run on these servers. The network security protection set up includes firewalls and intrusion detection systems to ensure the security of the entire platform. (3) The online experimental equipment group contains several kinds of equipment, which have together facilitated remote online joint experiments between multiple students.
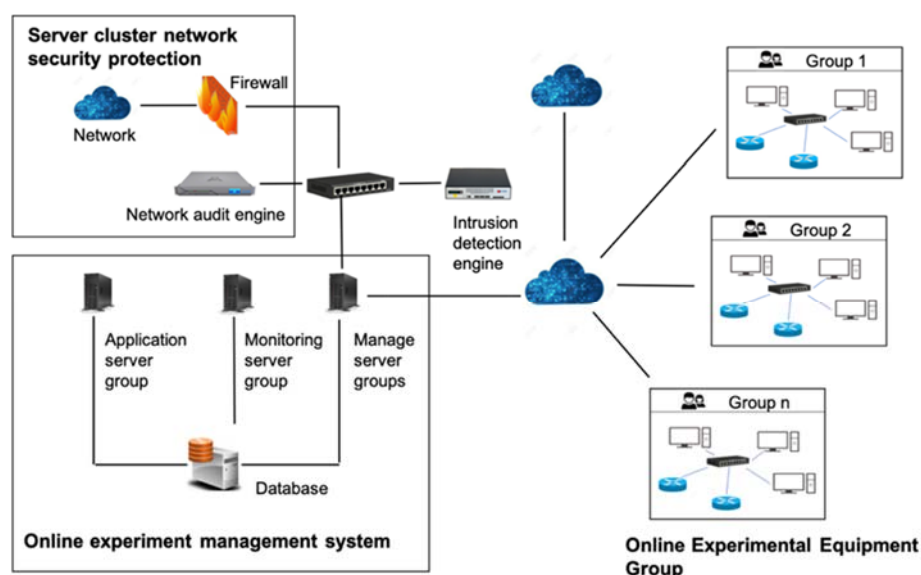

Figure 1. Overall structure diagram of the computer network online experiment platform.

On this experimental platform, students could complete most of the experiments that could normally only be performed in a network laboratory, including the basic configuration of network equipment and comprehensive network construction.

4. Computer Network Experiment Based on SPOC

Aiming at the requirements of network experiment teaching for undergraduates and postgraduates, we designed the network experiment content in a hierarchical fashion, and

the integrated teaching mode of this research was thus realized, as shown in Figure 2. (1) The basic principles of the network are the experimental basis, and this cultivates students' core understanding of and ability to apply network technology and network engineering practice. (2) Research experiments, including new network technology experiments that are more complex or difficult, mainly cultivate the understanding of and ability to apply network protocol and network engineering practice. (3) Innovative experiment and practice fund projects, focus on cultivating the network protocol design and development capabilities of a few high-achieving students, and guide students to design and implement new network protocols, algorithms, or systems.
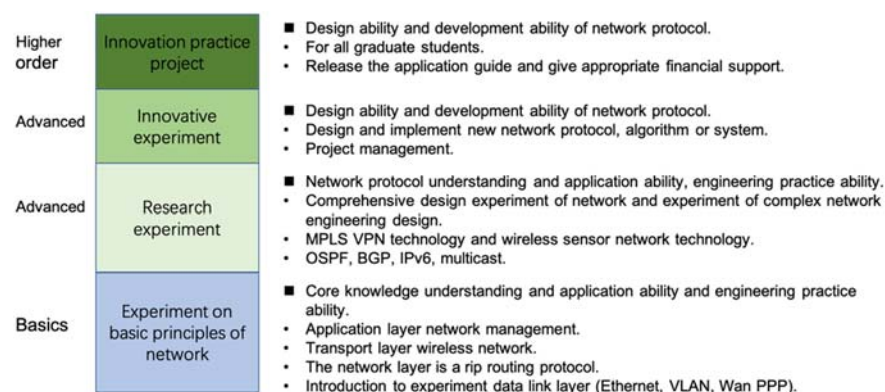

Figure 2. Multi-level integrated network experiment content of the research.

According to the integrated teaching mode of this research, we have designed the network experiment content of each chapter for both undergraduate and graduate students. Table 1 shows the network experiment courses we designed for undergraduates and graduate students. Here, the undergraduate course includes eight experiments and the postgraduate course includes seven experiments. The network experiment course includes learning videos, multiple-choice questions corresponding to the videos, and peer grading questions. In Table 1, RIP means routing information protocol, OSPF means open shortest path first, BGP means border gateway protocol, and TCP means transmission control protocol/internet protocol.

Table 1. Computer network experiment content.

Chapter ID	Undergraduate	Graduate
1	Getting started with network experiments	Comprehensive experiment of link layer and network layer
2	Data link layer experiment	OSPF routing protocol experiment
3	Network layer experiment	BGP routing protocol experiment
4	Transport layer experiment	TCP congestion control experiment
5	Application layer experiment	Network management experiment
6	RIP routing protocol experiment	IPv6 technology experiment
7	OSPF routing protocol experiment	Multicast technology experiment
8	Network engineering comprehensive experiment	—

4.1. Video Learning Courses Helpful for Experiment Preview

As shown in Figure 3, the video learning resources for each chapter of the SPOC course are screened by teachers in a targeted manner. Each section of each chapter contains a learning video; each video is 5–10 min in length, and the videos mainly include the key content that students need to master in the chapter, the areas that students often struggle to understand, and assessment of knowledge difficulty. Using these videos, students can preview the subject before class, determine the knowledge structure of each chapter using the textbook, and identify their own knowledge weaknesses.
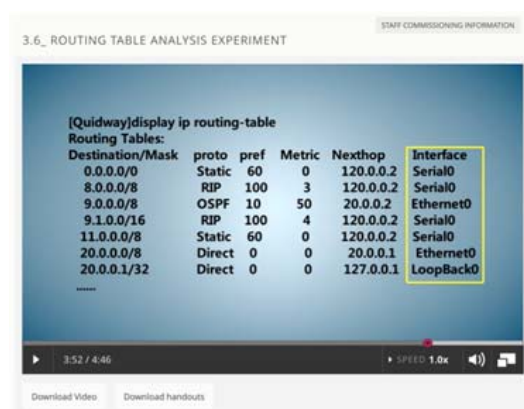
Figure 3. Video learning resources for network experiments.

4.2. Multiple Choice Exercises to Measure Students' Knowledge Mastery

After the learning videos, there are multiple choice questionnaires with 8 to 14 questions on each video. Each multiple choice problem corresponds to the subject of the chapter. Students can use this to assess their knowledge after the preview, as shown in Figure 4. We have established four or more answer categories for each question, and each problem necessitates a certain degree of thinking. When students have completed each multiple-choice question, they can click the button "CHECK THE ANSWERS"; the system will give a hint as to whether the student's answer is correct, but will not directly give the correct answer. For each problem, the student has three opportunities to click "CHECK THE ANSWERS". When students are not familiar with the topic, they can click "SHOW ANSWERS" after clicking "SAVE" to check where their knowledge is lacking.

Figure 4. Multiple-choice questions in the network experiment course.

4.3. Peer Grading for Students' Comprehensive Ability

Figure 5 shows an example of the BGP design and implementation in this course. Figure 5a gives the title description of the assignment, which requires students to design and configure routing strategies. Figure 5b shows the connection diagram of routers and switches. Students need to deploy routers in the actual physical network environment and add routing configurations as needed. The assignment includes five questions, concerning designing a network diagram, conceiving of subnetting, basic BGP configuration and routing strategy configuration, and experimental verification. Each problem is designed to check the student's mastery of specific skills in the computer network laboratory. For example, problem 2 checks whether students can correctly divide the network into multiple subnets and assign host addresses to each. Problem 3 examines students' mastery of the border gateway protocol, and problem 4 evaluates students' configuration of routers and switches and their static routing capabilities.
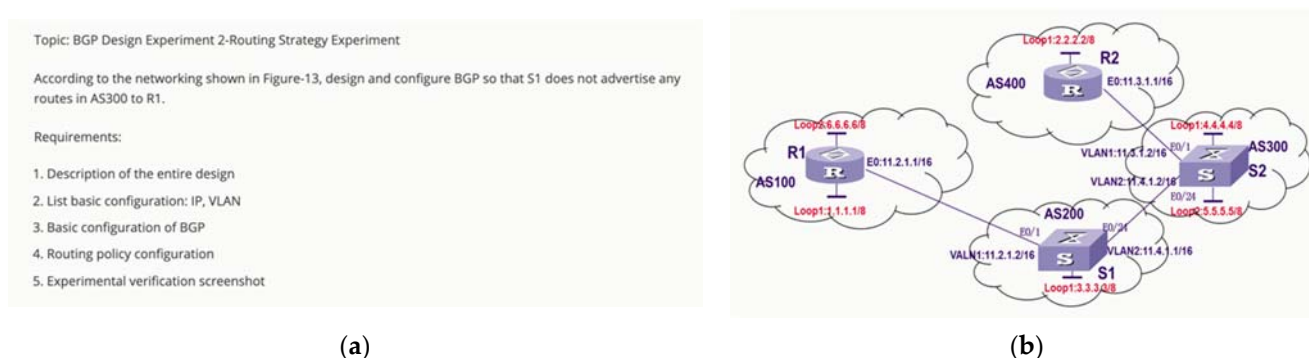

(**a**) 　　　　　　　　　　　　　　　　　　　(**b**)

Figure 5. BGP design experiment: (**a**) describes the background and requirements of the assignment, and lists five questions that students need to answer; (**b**) shows the network topology of this assignment.

In the process of peer grading, graders will evaluate the answers of their peers to each question according to the grading rules specified by the TAs. Figure 6 shows the grading for problem 1. The grading provides four categories (completely correct, mostly correct, partly correct, and wrong). In order to make it easier to analyze the behavior of the grader, we set each category to contain only one knowledge point. The grading rules instruct the student to grade each answer based on their subjective judgment of its quality. The grading rules specify an answer category and corresponding grade for each problem.

Figure 6. Types of questions and grades for each category.

5. Online Blended Learning Based on Network Experiment Courses

The online blended learning proposed in this paper refers to students' pre-class preparation, after-class or online experiments, and peer grading. Pre-class preparation requires students to familiarize themselves with the subject at hand using the textbook and the course video, and then complete the exercises at the end of each chapter to reach preliminary familiarity with the experimental content and procedures. After-class or online experiments require students to make an appointment during class time and complete the experiment within the specified time. Since many students need to perform experiments

within a certain period of time, they can also make an appointment for online experiments and thus complete network experiments outside of class. After students have completed the network experiment, we will ask them to submit their own experimental process and results to the SPOC system in the form of a submission. The system will redistribute all students' submissions to different students, and peer review thus takes place at the same time. Peer grading is an investigation of a student's comprehensive ability. Through peer grading, students can not only identify and solve the problems in their work, but also further deepen their experimental knowledge. In order to improve the online blended learning, we have also addressed four problems in peer grading: the evaluation of students' cognitive ability, task assignment in peer grading, removing the noise data in the peer grades, and improving students' conscientiousness.

5.1. Necessity of Peer Grading and Basic Environment

We have designed an SPOC environment, which contains a relatively small number of students compared with massive open online courses (MOOCs). Although there are differences between SPOCs and MOOCs, SPOCs involve the same peer grading approach as MOOCs. This is because teachers must review hundreds of open-ended essays and exercises, such as mathematical proofs and engineering design questions, at the end of the final semester. Considering the scale of the submissions, if teachers have to check all the submitted content alone, they may become overwhelmed and not have enough time and energy to carry out other teaching tasks.

In order to assign a reasonable grade to open coursework, some researchers have proposed the use of an autograder, but an autograder also has its own shortcomings: (1) compared with the feedback provided via manual review, automatic grading usually only assesses whether the answer is right or wrong, and cannot give constructive feedback; (2) it is also unable to judge open and subjective topics, such as experimental reports and research papers.

Therefore, it is best to use peer grading for these kinds of open problems. Peer grading means that students are both the submitters and reviewers of the assignments—one submission is reviewed by multiple graders, and one grader reviews multiple submissions. Our network experiment course is designed such that one submission is reviewed by four graders, and one grader reviews four submissions.

In establishing peer grading, we solved four problems, including the reliability estimation of the grader, the assignment of tasks, the issue of human-machine integration, and peer prediction.

5.2. Grade Aggregation Model Based on Ability Assessment

(1) Problems in peer grading

In the grade aggregation model, we must consider the reliability of the grade in order to improve the accuracy of grade aggregation. However, in the SPOC environment, the learners know each other and can communicate arbitrarily, and there is often collusion when reviewing submissions, meaning the law of large numbers cannot be used to model reliability characteristics. Therefore, accurately modeling the reliability characteristics of the graders is an important problem we have to solve.

(2) Personalized ability estimation model

We assess student abilities using the historical data of their watching videos and performing exercises as the basis for a reliability evaluation in order to establish a two-stage personalized ability evaluation model.

First, we extract some characteristic values related to the students when they watch the video, including the total time for which the students watch the video, the number of times the video is forwarded, the number of videos watched, the length of forum posts, the total length of videos watched, the number of pauses, and the number of rewinds. Then we

use the logistic regression method to fit the above features to the equation representative of the student's initial ability value. First, we define Equation (1):

$$p(y = 1 | X; W) = g\left(w^T X\right) = \frac{1}{1 + e^{-W^T X}} \tag{1}$$

Here, W represents the weight of each feature, and X represents the value of each feature, $y \in \{0, 1\}$ represents whether the feature exists. The maximum likelihood estimation method can be used to solve the parameter W, the iteration process is Equation (2), and the α is the expansion factor,

$$WW - \alpha \sum_{i=1}^{n} \left(y_i - g\left(w^T X_i\right)\right) X_i \tag{2}$$

In the second stage, we enter each student's answer sequence into the knowledge tracking model to evaluate their learning. As shown in Figure 7, the ability of each student is predicted using the output of the first stage as the S-node training personalized knowledge tracking model. In Figure 7, node S is a student node, node K is a knowledge level node, and node Q is a question answer node. $P(L_0)$ represents the initial probability of knowledge mastery, $P(G)$ represents the probability of a correct guess, $P(S)$ represents the probability of error, and $P(T|S)$ represents the probability of personalized learning.
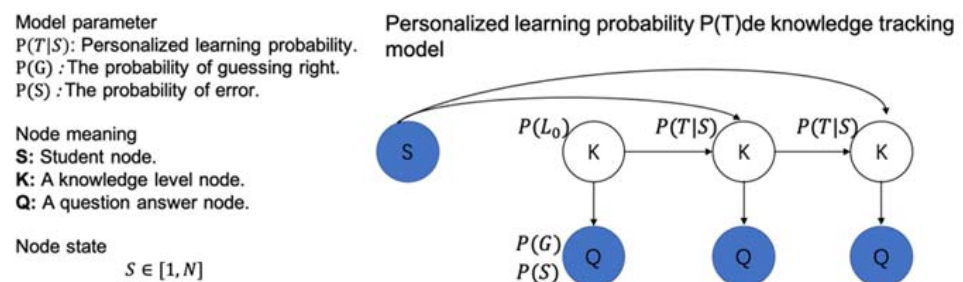

Figure 7. KT model of personalized learning probability.

(3) Grade aggregation model

In this work, we remodeled its reliability characteristics based on the ability of the graders. The first model we designed here is called PG6. Here, the reliability variable τ_v follows a gamma distribution, wherein the hyperparameters β_0, η_0, μ_0 and γ_0 are empirical values, and a_v is the shape parameter, the value of which represents the output result of the two-stage model we designed.

$$\tau_v \sim \mathcal{G}(a_v, \beta_0) \tag{3}$$

$$b_v \sim \mathcal{N}(0, 1/\eta_0) \tag{4}$$

$$s_u \sim \mathcal{N}(\mu_0, 1/\gamma_0) \tag{5}$$

$$z_u^v \sim \mathcal{N}(s_u + b_v, 1/\tau_v) \tag{6}$$

where

a_v represents the previous evaluation ability of each student grader (v).

τ_v represents the reliability distribution of the grader v.

b_v represents the deviation of the grader v.

s_u represents the true grade of submission (u).

z_u^v represents the grade given by the grader v to the submission u.

We also designed the PG7 model, which is similar to the PG6 model and in which the reliability variable τ_v follows a Gaussian distribution. The mean parameter in the Gaussian distribution is a_v.

$$\tau_v \sim \mathcal{N}(a_v, 1/\beta_0) \tag{7}$$

High-scoring students are believed to perform peer grading tasks more reliably. Based on this assumption, we can further extend the PG7 model into the PG8 model, wherein the hyperparameters θ_1 and θ_2 are empirical values.

$$\tau_v \sim \mathcal{N}(\theta_1 a_v + \theta_2 S_v, 1/\beta_0) \tag{8}$$

5.3. Assignment of Tasks in Peer Grading

(1) Task assignment in peer grading

In peer grading, unreasonable task assignment has a great influence on the aggregation of grades. In peer grading, each student plays two roles, as both a submission submitter and a submission reviewer. The essence of the grading task assignment problem is how to form a grading group. Many online experiment platforms use simple random strategies and do not consider differences between students. In fact, different students have different knowledge levels, so the reliability of grading also differs. Random assignment may result in some grading groups containing students with better knowledge, while other groups are composed of students with lower levels of knowledge, which thus affects the overall accuracy of homework grading.

(2) MLPT-based task assignment model

There are many similarities between the grading submission assignment problem and the task scheduling problem, so the LPT algorithm also provides a good solution for the former. The problem of assignment can be understood from another perspective. Assigning submissions to students for review is equivalent to assigning students to assignments for review, so the student set $V = \{v_1, v_2, v_3 \ldots v_n\}$ can be regarded as the set of tasks to be assigned in the scheduling problem $J = \{j_1, j_2, j_3 \ldots j_n\}$. The knowledge level $w(v_i)$ of each student is understood as the processing time $t(j_i)$ required by the task, and the submission set $U = \{u_1, u_2, u_3 \ldots u_n\}$ is understood as the set of processors $M = \{m_1, m_2, m_3 \ldots m_p\}$. Therefore, in the task scheduling problem, the tasks are assigned to the machines to ensure that the running times of each are as similar as possible, which is equivalent to assigning students to submissions in such a way as to ensure that the knowledge levels of the reviewing students are as similar as possible.

By integrating the similarities and differences of submission assignment and task scheduling, we propose an MLPT (modified LPT) algorithm to solve the submission assignment problem in peer grading. Its pseudo code is shown in Algorithm 1.

Algorithm 1. MLPT submission assignment algorithm.

MLPT task assignment algorithm

1: Sort the graders according to their knowledge level so that $w(v_1) \geq w(v_2) \geq \ldots \geq w(v_n)$

2: For $i = 1, 2 \ldots, n$:

A collection $U(i)$ of assignments reviewed by students v_i $U(i)$ is empty, $U(i) \leftarrow \varnothing$

3: For $j = 1, 2 \ldots, n$:

The comprehensive level L_j of the grading group corresponding to submission u_j is assigned 0, $L_j \leftarrow 0$

Assignment u_j corresponds to the student set $V(j)$ of the grading group to be empty, $V(j) \leftarrow \varnothing$

4: For $i = 1, 2, \ldots, n$:

For $k = 1, 2, \ldots, m$:

In the assignment set $\{u_j | j \neq i\}$, select the assignment with the smallest comprehensive level of the corresponding grading group, and mark it as u_j,

Add current student v_i to the grading group corresponding to homework u_j, $V(j) \leftarrow V(j) \cup v_i$,

Update the comprehensive level L_j of the grading group corresponding to the assignment u_j, $L_j \leftarrow L_j + w(v_i)$,

Add submission u_j to the student set $U(i)$ reviewed by student v_i, $U(i) \leftarrow U(i) \cup u_j$.

5.4. Peer Grading Model Based on Human-Machine Hybrid Framework

(1) The problem of a large number of random grades in peer grading

The probability graph model that is suitable for peer grading in MOOCs may not perform well in SPOCs. Because students know each other in an SPOC environment, they often assign a random grade to a given assignment, rather than reviewing it seriously. In fact, we found that some graders simply give the submission a full mark or a high grade in order to complete the task quickly. This kind of careless review behavior prevents the basic assumptions in the Bayesian statistical model from being met, and will inevitably produce data noise that reduces the performance of the model.

(2) A framework for peer grading of tasks between human and machine

In order to solve the problem of data noise in the grade aggregation model, we propose a framework based on a combination of human and machine. Figure 8 shows the human-machine hybrid framework, which consists of three main parts: the task autograder, the grade filter, and the PG model.
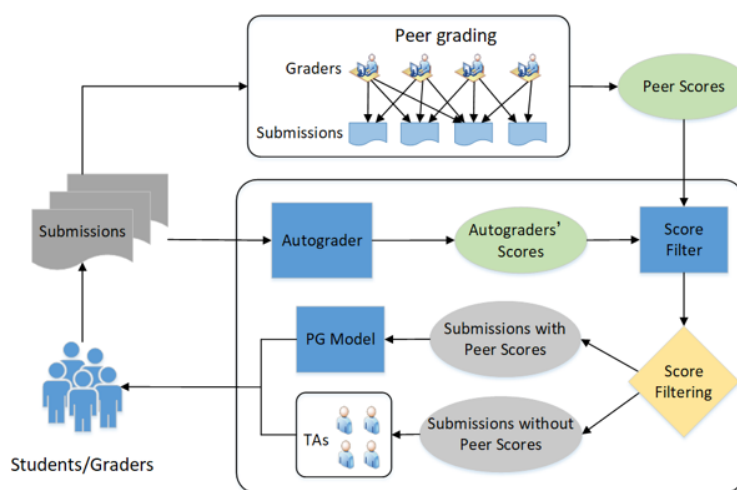

Figure 8. Human-machine hybrid submission evaluation framework.

In the peer grading process, the system first assigns each grader a peer grading task. When the grader reviews a submission, the autograder will review the same submission at the same time. The result of the reviewer is passed to the grade filter. The grade filter compares the grades of the autograder and the grader. If the deviation between the two grades exceeds a predefined threshold, the grader's grade is discarded. Under the coordination of the autograder and the grade filter, the framework divides the peer grades into two categories: a set of submissions contains the peer grades submitted by serious reviewers, which can be aggregated by the PG model to extrapolate the final submission grade; another group of student submissions without peer grades will be sent to the TAs for peer grading. In this way, we can resolve the problem caused by random peer grading, thereby improving the performance of the PG model.

5.5. Eliciting Truthfulness with Machine Learning

(1) The problem of a non-serious attitude to review in the peer grading

The degree of seriousness of the graders during the review directly affects the accuracy of the grade aggregated after the peer grading. Through statistical analyses of the peer grades from 2015 to 2018 and comparisons with groundtruth, we found that the number of peer grades given by graders is generally higher than the groundtruth. Therefore, we concluded that many students did not participate in peer grading. This may be because they were trying to cope with the task; grades may also have been deliberately given indiscriminately, or they may have been unclear but the reviewer was unwilling to think about it seriously, so they gave inaccurate grades.

(2) Reward and incentive mechanism model

In order to completely eliminate chaotic scoring behavior and improve the motivation of students in peer grading activities, we designed a reward and incentive mechanism to increase the seriousness of the graders. Here, the review result of the grader is compared with that of the autograder, and the grader is rewarded or punished according to the adherence to uniform distribution. This means that if the grader's result seems incorrect, they can only reevaluate if they want to give the correct result. In this way, the autograder can motivate the grader to review more carefully. At the same time, when the grader's result is inconsistent with that of the autograder, if they want the maximum reward, they must give an accurate review.

Here we define $\mathrm{p}(s_{auto}^r)$ as the accuracy of autograder, and the probability that autograder chooses A, B, C, and D is equal, that is,

$$\mathrm{p}(A) = \mathrm{p}(B) = \mathrm{p}(C) = \mathrm{p}(D) \tag{9}$$

Equation (9) is the previously assessed probability of a topic being evaluated correctly, and $\mathrm{p}(a|A)$ is the probability of the grader i reporting a when the prior value is A, assuming that $f(a|A)$, $f(a|B)$, $f(a|C)$, and $f(a|D)$, the probability of the grader i choosing a is:

$$\mathrm{Pr}(s_i^a) = f(a|A)p(A) + f(a|B)p(B) + f(a|C)p(C) + f(a|D)p(D) \tag{10}$$

Thus the probability of the autograder's review result being a is:

$$g(s_{auto}^a|s_i^a) = f(a|A)\frac{f(a|A)p(A)}{\mathrm{Pr}(s_i^a)} + f(a|B)\frac{f(a|B)p(B)}{\mathrm{Pr}(s_i^a)} + f(a|C)\frac{f(a|C)p(C)}{\mathrm{Pr}(s_i^a)} + f(a|D)\frac{f(a|D)p(D)}{\mathrm{Pr}(s_i^a)} \tag{11}$$

The result calculated from Equation (11) is the probability of the autograder obtaining a, which is converted from a probability to a grade. We use the logarithmic grading rule, as shown in Equation (12):

$$\mathrm{score}(s_{auto}^n|re_i) = \ln(g(s_{auto}^n|re_i)) \tag{12}$$

Here, n represents the number of categories of answers in a question. We use the report given by grader i to predict the grade value of the autograder. Following the above calculation of the autograder's grade value, Equation (11) is used to predict the reward for grader i:

$$\sum_{n=1}^{4} \mathrm{score}(s_{auto}^n|re_i)g(s_{auto}^n|s_i) \tag{13}$$

The only way to maximize Equation (13) is to make $re_i = s_i$. Because the autograder only gives its own review results, if the grader i wants to get the maximum reward, the report they give must be consistent with what they really think.

6. The Learning Effect of Online Blended Learning

Our online blended learning method of computer network experimentation improves the learning effect, optimizing the teaching mode of the network experiment, improving experimental knowledge through peer grading, and clearly improving exam grades.

6.1. Optimize the Teaching Mode of Network Experiment

In the traditional teaching mode, students can only go to the network laboratory to undertake experimental learning, and their learning is inevitably restricted by time and space. A considerable part of the experimental time is occupied by traditional verification and analytical experiments, and time and conditions are insufficient to carry out more difficult research and innovation experiments. Via the online learning model, it is easy to carry out experiments of medium or lower difficulty as online experiments. As part of the preparation process, students are required to complete the experiment online before participating in the experiment in class. Teachers and TAs provide remote online experiment

guidance. After students have completed the basic online experiments, they can conduct comprehensive analytical and design experiments in the laboratory, and can even design experimental content independently. The classroom is mainly for one-to-one guidance and small-scale discussions.

By offering pre-experiment explanations and experimental previews in a video format, the online blended learning method combining online experiments, offline experiments, and peer grading greatly improves the teaching effects, with students able to complete basic online experiments later. Comprehensive design experiments with a certain degree of difficulty are directly carried out in the laboratory. This not only broadens the content of experimental teaching, but also increases its depth and difficulty.

6.2. The Promotion of Peer Grading to the Benefit of Network Experimental Teaching

By solving the four problems in peer grading: the evaluation of students' cognitive ability, task assignment in peer grading, removing the noise data in the peer grades, and improving students' conscientiousness, the online SPOC course helps students to gain a deep understanding of network experimentation, and while strengthening their hands-on ability, it also helps students to understand the inner principles of the experiment.

6.2.1. Experimental Results of the Grade Aggregation Model Based on Ability Assessment

In this paper, we compare the results of our PG6-PG8 model with those of a simple median model based on the baseline, as well as with the PG1-PG3 models proposed in [7] and the PG4-PG5 models defined in [8]. The evaluation index is the root mean square error (RMSE), which is calculated via the deviation between the ground truth grade given by the TAs and the peer grade given by the student grader. In the review of students' abilities, since the first and last problems in peer grading are related to visualization, in order to make the experiment simple, we only assessed the second, third, and fourth questions. The experimental results shown in Tables 2 and 3 are the average RMSE values of these models over 10 experiments.

Table 2. The experimental results for the peer grades given by the graders who did not make submissions, without filtering.

Model		Problem 2		Problem 3		Problem 4	
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Baseline	Median	17.65	—	13.75	—	13.23	
Piech et al. 2013 [7]	PG1	14.43	0.02	12.76	0.04	11.69	0.02
	PG3	14.25	0.03	12.43	0.03	11.54	0.04
Fei Mi et al. 2014 [8]	PG4	14.28	0.02	12.46	0.06	11.32	0.06
	PG5	14.35	0.03	12.42	0.04	11.30	0.07
	PG6	14.33	0.09	12.45	0.04	11.35	0.05
Models based on ability	PG7	14.36	0.08	12.43	0.05	11.31	0.06
	PG8	14.33	0.09	12.39	0.06	11.27	0.05

Table 3. The results of the experiment assessing the initial ability of a grader who did not offer a submission.

Model		Problem 2		Problem 3		Problem 4	
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Baseline	Median	1822	—	14.41	—	14.21	
Piech et al. 2013 [7]	PG1	15.63	0.02	12.56	0.04	12.40	0.03
	PG3	15.58	0.01	12.37	0.05	12.33	0.01
Fei Mi et al. 2014 [8]	PG4	15.67	0.08	12.40	0.07	12.35	0.08
	PG5	15.58	0.07	12.33	0.09	12.32	0.05
	PG6	15.58	0.01	12.35	0.06	12.36	0.04
Models based on ability	PG7	15.60	0.04	12.37	0.07	12.38	0.05
	PG8	15.50	0.08	12.29	0.09	12.35	0.09

In our experiments, we found that if the student graders who did not submit submissions were given random grades depending on their previously assessed ability, the fluctuations in predicted grades would increase, because if there is no accurate previous estimate of a students' ability, it is difficult to give a reasonable estimated grade for the submission. It can be seen from Tables 2 and 3 that the average performance of PG8 is the best among all the models. The best results can be obtained by combining the grader's grading ability and the real grade with appropriate weights. Therefore, the PG8 model is the best approach to aggregating the peer grades in SPOCs. This section aggregates the peer grades based on this ability assessment method, then evaluates the effect of students participating in peer grading, provides reasonable grades for the experimental online submissions of students, and in so doing reduces the teachers' burden.

6.2.2. The Results of the Task Assignment Experiment for Peer Grading

Figure 9 shows the experimental results of using the MLPT algorithm for peer grading. The ordinate represents the reduction in the variance of the grading group after replacing the random algorithm with the MLPT algorithm. It can be seen that the MLPT algorithm is effective in reducing the variance in the grading group level, especially when the number of participating students is small and the redundancy is small, under which conditions the advantages of the MLPT algorithm are more prominent. The MLPT algorithm solves the problem of unreasonable task allocation in peer grading.
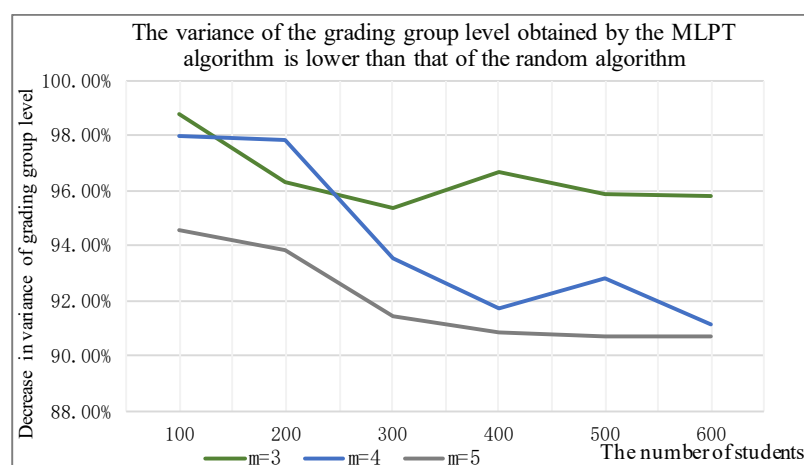

Figure 9. Comparison of experimental results between the MLPT algorithm and the random algorithm.

6.2.3. The Results of a Human-Machine Hybrid Framework for Peer Grading

In this experiment, the RMSE calculation result is calculated from the average value of 500 iterations. It can be seen from Table 4 that no matter which PG model is used, the human-machine combination framework can achieve the best performance. Compared with the original solution without fraction filtering, the average RMSE is reduced by 8–9 points. This result confirms that in the presence of random review behavior, the human-machine combined framework model can improve the prediction accuracy of the PG model. Using this human-machine hybrid framework, the problem of noisy data in grade aggregation is solved, and the accuracy of the grade aggregation model is improved.

Table 4. Experimental results of the human-machine framework.

Model	RMSE				
	Initial Data Set without Filtering	Original Data Set Based on Naive Bayes Autograder	Data Set with Random Filtering	Data Set with Anomaly Detection	CNN-Based Autograder Dataset
PG1	21.34	21.90	21.05	16.89	12.54
PG3	20.12	22.02	21.10	16.31	11.43
PG4	21.90	22.43	22.30	17.20	11.32
PG5	21.65	21.67	20.50	17.30	12.20

6.2.4. The Results of Eliciting Truthfulness with Machine Learning

(1) Comparison of error statistics experiment results with and without autograder.

We compared the results of a peer grading process using an autograder and the same with no autograder in the BGP experiment of the network experiment class, and we found that the autograder can motivate the graders to review more carefully. In Figure 10, we set the number of students participating in the experiment, the number of peer grades, and the interval between the horizontal and vertical axes in the histogram in the two graphs to be equal. It can be seen from the figure that after adding an autograder as the object of the peer prediction, the number of graders with larger error values is significantly reduced, and especially those with an error value of 35–50. The resulting peer grades contain very little noise, so they can be effectively applied to the PG model.
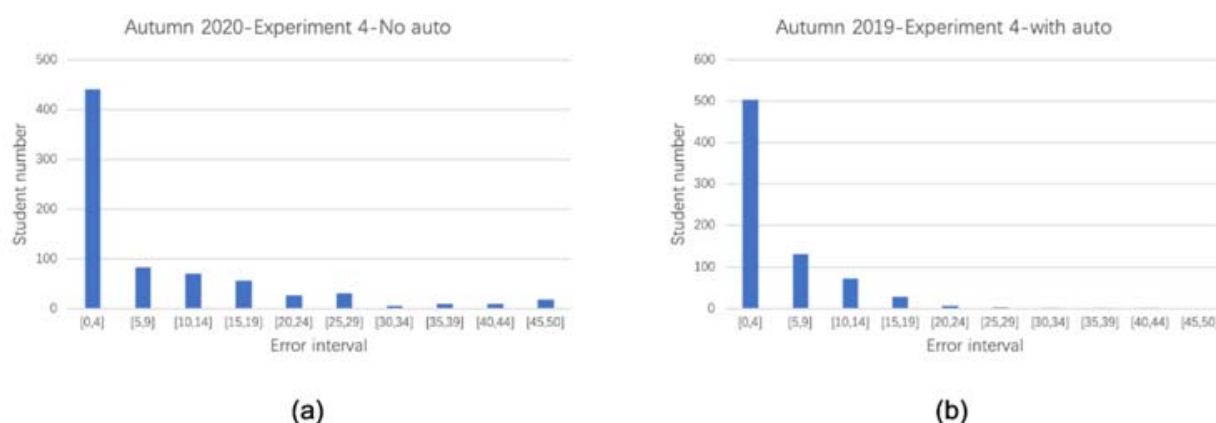

Figure 10. Comparison experiment with and without autograder. (**a**) We added autograder to the peer grading in the autumn of 2020. (**b**) We did not add autograder in the peer grading in the autumn of 2019.

(2) Comparison of RMSE when autograder is added and when it is not added

In Table 5, the error grade is taken as the absolute value of the difference between the grader's peer grade and the final real grade for the same submission. The purpose of the RMSE in the table is to measure the degree of dispersion of the data. It can be seen from the table that an autograder was not added to experiment 4 in the fall of 2020, while it was added in the autumn of 2019. The error mean and standard deviation of the latter both decreased, and the RMSE value decreased by 6.39 points, or 82.77%. The data show that after adding an autograder, the gap between the peer grade and the actual grade is significantly reduced, which means that the quality of student peer grading is significantly improved. By eliciting truthfulness via machine-learning in the peer grading process, the seriousness with which students take peer grade is effectively improved.

Table 5. Experimental results of the game process with and without autograder.

	Autograder	Number of Submissions	Number of Grades	Mean	Standard	RMSE
Autumn—2019	yes	198	750	4.91	5.72	7.54
Autumn—2020	no	198	750	8.87	10.97	14.11

6.3. Verify the Improvement of the Effect of Network Experiment Teaching

In the context of the computer network online experiment platform, by optimizing the teaching mode of the network experiment, and enhancing the depth and breadth of the experimental learning content through online blended learning, the students' experimental learning can be improved, as can the teaching effect of the computer network experiment.

We compared the results of an online experiment exam after the online experiment platform was employed in actual teaching in 2015 with the results from previous years. (Figure 11).
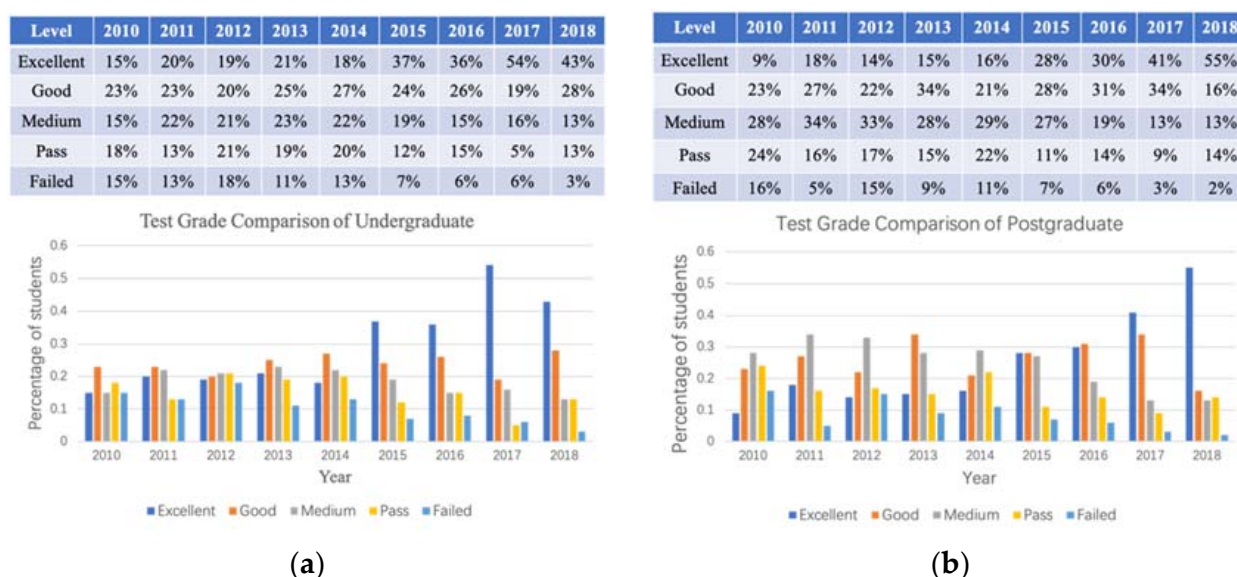

Figure 11. Comparison of exam grades. (**a**) Statistical comparison of undergraduates' experimental exam grades. (**b**) Statistical comparison of postgraduates' experimental exam grades.

Through analyses and comparisons of the network experimental teaching data from 2010 to 2018, we see that the use of online blended learning methods offers a wider range of experimental teaching content, more depth, greater difficulty, better teaching effects, and significant increases in the percentage of excellent and good results. The "fail" percentage dropped significantly. Through the analysis of students' final exam grades, we see that blended learning methods can effectively improve students' learning.

7. Conclusions

The online blended learning method combining computer network experiments and online SPOC course learning enables students to carry out hardware-based online experimental learning anytime and anywhere, and effectively improves the teaching level and learning effect. It also addresses the difficulty in making appointments and the spatial and temporal limits for students when performing experiments, and provides students with a free and flexible experimental environment; that is, they can exercise their hands-on ability and deepen their understanding of theoretical knowledge.

Moreover, the online blended learning method proposed in this paper not only enables students to conduct online experiments and learn in an actual teaching environment; it also

enables us to analyze the learning trajectories of students on experimental courses, analyze their mastery of knowledge points, and provide them with targeted guidance. For example, fine-grained learning behavior data can also be obtained. Through continuous quantitative analyses of micro-teaching activity data, the cultivation of students' professional abilities can be tracked. This can effectively improve the quality of teaching.

The blended learning method not only provides students with flexible and diverse learning, but also effectively promotes the improvement of teaching quality, which facilitates analysis and research on the relationship between various elements within the applied sciences.

The blended learning method proposed in this paper can also be applied to other courses by converting the traditional teaching environment to online teaching, and analyzing the learning behaviors of students. However, some problems may also be encountered in the specific cases of implementation, such as how to evaluate the learning more effectively. To solve these problems, we require further literature and material studies.

Author Contributions: This paper was completed by a number of authors, each of whom contributed: Conceptualization, Y.H.; methodology, W.W.; software, Y.L.; data curation, L.Z.; writing—original draft preparation, Y.H.; writing—review and editing, Y.H. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this paper is from the student behavior data of Beihang University. Its website is http://mooc.buaa.edu.cn/dashboard, accessed on 1 July 2021.

Conflicts of Interest: The authors declare that they have no competing type of interests.

References

1. Bettens, A.M.; Morrell, B.; Coen, M.; McHenry, N.; Wu, X.; Gibbens, P.; Chamitoff, G. UnrealNavigation: Simulation software for testing SLAM in virtual reality. In Proceedings of the AIAA Scitech 2020 Forum, Orlando, FL, USA, 6–10 January 2020; 2020; p. 1343.
2. Staubitz, T.; Meinel, C. Collaborative Learning in SPOCs Approaches and Experiments. In Proceedings of the 2018 IEEE Frontiers in Education Conference (FIE), San Jose, CA, USA, 3–6 October 2018; pp. 1–9.
3. Mhlanga, D.; Moloi, T. COVID-19 and the Digital Transformation of Education: What Are We Learning on 4IR in South Africa? *Educ. Sci.* **2020**, *10*, 180. [CrossRef]
4. McGrath, G.M.; Lockstone-Binney, L.; Ong, F.; Wilson-Evered, E.; Blaer, M.; Whitelaw, P. Teaching sustainability in tourism education: A teaching simulation. *J. Sustain. Tour.* **2020**, *29*, 1–18. [CrossRef]
5. Dhawan, S. Online learning: A panacea in the time of COVID-19 crisis. *J. Educ. Technol. Syst.* **2020**, *49*, 5–22. [CrossRef]
6. Hui-Ling, H.O.U. Exploration on Experiment Teaching Reform Based on Computer Network Teaching Platform. *Comput. Telecommun.* **2020**, *1*, 4–6.
7. Dziabenko, O.; Adorno, D.P. Application of remote experiments in a secondary school using SPOC approach. In Proceedings of the 2017 4th Experiment@ International Conference (exp. at'17), lgarve, Portugal, 6–8 June 2017; pp. 191–195.
8. Nedungadi, P.; Prabhakaran, M.; Raman, R. Benefits of Activity Based Learning Pedagogy with Online Labs (OLabs). In Proceedings of the 2017 5th IEEE International Conference on SPOCs, Innovation and Technology in Education (MITE), Bangalore, India, 27–28 October 2017; pp. 52–56.
9. Noviana, M.; Sukardi, S.; Suryanti, N.M.N. Learning Process during Covid-19 Pandemic from Various Variables in Senior High School. *SAR J.* **2020**, *3*, 160–165. [CrossRef]
10. Husár, J.; Dupláková, D. Evaluation of foreign languages teaching in LMS conditions by facility and discrimination index. *TEM J.* **2016**, *5*, 44–49.
11. Piech, C.; Huang, J.; Chen, Z.; Do, C.; Ng, A.; Kolle, D. Tuned Models of Peer Assessment in SPOCs. *Comput. Sci.* **2013**, arXiv:1307.2579.

12. Mi, F.; Yeung, D.Y. Probabilistic Graphical Models for Boosting Cardinal and Ordinal Peer Grading in SPOCs. In Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015; pp. 454–460.

13. Yan, Y.; Rosales, R.; Fung, G.; Dy, J.G. Active learning from crowds. In Proceedings of the 28th international conference on machine learning (ICML-11), Bellevue, WA, USA, 28 June–2 July 2011; pp. 1161–1168.

14. Gao, A.; Wright, J.R.; Leyton-Brown, K. Incentivizing evaluation via limited access to ground truth: Peer-prediction makes things worse. *arXiv* **2016**, arXiv:1606.07042.

15. Mita', D.; Dupláková, D.; Duplák, J.; Mita'ová, Z.; Radchenko, S. Implementation of Industry 4.0 Using E-learning and M-learning Approaches in Technically-Oriented Education. *TEM J.* **2021**, *10*, 368–375. [CrossRef]

16. Chengbin, Q.; Yongqiang, C.; Shanshan, L.; Youjian, Z. Exploration of the Computer Hardware Experiment teaching method based on the cloud platform. In Proceedings of the 2016 IEEE Frontiers in Education Conference (FIE), Erie, PA, USA, 12–15 October 2016.