

Good Performance Factor Analysis of Online Learning Case Data*

Yanxia Pang Jiameng Cai Zhen Xiang

School of Computer and Information Engineering
Shanghai Polytechnic University, Shanghai, China
yxpang@sspu.edu.cn, 18317028728@163.com,
1770427997@qq.com

Yan Pan

Shanghai Foreign Language
School Affiliated to SISU
Shanghai, China
lovymao@163.com

Abstract—Online learning is widely applied especially in the pandemic period. However, online learning performance varies in a wide range. Good performance leads to satisfying learning experience and less drop-out. What contributes to the good performance? Based on the learning data, related factors are analyzed in this paper. A self-paced online course with big learning data from a Stanford platform is selected as a case for this research. Related surrounding factors cover location, language, time, date, semester, etc. The difference of learning event numbers between learners is also found great. Research on the data of good performance learners tries to find out the correlations between study habits and good performance. First, a good performance learner with the richest data is picked out as a sample. Then, the findings of the sample analysis are tested on the data of another 4 good performance learners. Fixed learning time mode, effective video viewing and balanced effort distribution over various learning resource are found the good performance learners' common features.

Keywords—online learning, data analysis, good performance, learning event

I. INTRODUCTION

With the development of technology, online learning is widely accepted, especially in the period of pandemic. Quarantine is suggested to interrupt the virus spreading. Longhurst thinks the COVID19 pandemic leads to an almost immediate switch to distance learning by higher education institutions [1]. Eom suggests online education can be a superior mode of instruction if it is targeted to learners with specific learning styles [2]. Individual differences have been recognized as an important factor in the learning process [3]. Baber analyzes the determinants of students' perceived learning outcome and satisfaction in online learning during the pandemic of COVID19 [4]. How to improve the performance of online learners? It is a problem deserving detailed analysis. This paper researches the correlations between learning performance and the factors including surroundings and study habits.

Learning style and study habits both influence the learning performance [5]. The research in [6] finds there is a systemic change in learning quality, and it evaluates the teaching and learning support of the network surroundings. Kuama

examines online language learning strategies used and affection in online learning of successful and unsuccessful online language students [7]. But the work is about language learning and ignores the surroundings and study habit factors. Most related work is based on the data from the questionnaires. Values from questionnaires depend more on the self-evaluation of learners. Online learning data recorded by the platform is more objective.

Learning analytics research has demonstrated that the large volume of learning data collected by modern e-learning systems could be used to recognize learner behavior patterns and could be used to connect these patterns with measures of student performance. [3] To make sure no interfering of the lecturers or platforms, a self-paced philosophy course is adopted as a case for this research which is provided by the Center for Advanced Research through Online Learning(<https://iriss.stanford.edu>).

II. THE ADOPTED CASE DATA

The learning data of the course is from Stanford's platform using the OpenEdX. 207 videos and 467 learning resource are included. 2076568 enrollments are recorded. 1048574 learners from 157 countries enrolled in the course learning. The data is big enough to avoid survivorship bias.

Learning resource analysis helps to find out what contributes to the performance. Pereira uses learning resource analysis as a tool for verifying the learning [8]. The learning resource of the data set is presented in different forms. It covers more than course, chapter, LTI, sequential, vertical, video and problem as Fig.1 lists. Their ratio comparison is listed. They count 47026, 58541, 1098, 86562, 556, 66361 and 232227 respectively. Learners interactive with different parts of the course and answer problems. The problems can be answered more than once, so the learner can achieve better scores. learners want to achieve good learning performance. Good performance leads to more learning effort devotion.

III. PERFORMANCE DIFFERENCE BETWEEN LEARNERS

The distribution of max grade is also demonstrated as Fig.2 shows. The grade values are in the range from 0 to 10. Data of learners with 0 max grade is filtered out since some learners may prefer watching videos to taking tests. 5 and 10 are mostly found as the max grade. There are obvious gaps

*Funding: This paper is supported by the first-class undergraduate profession building of Shanghai under Grant No. A30NH221903-0313.

between score 10, 5 and 1. The performance difference between learners is significant.

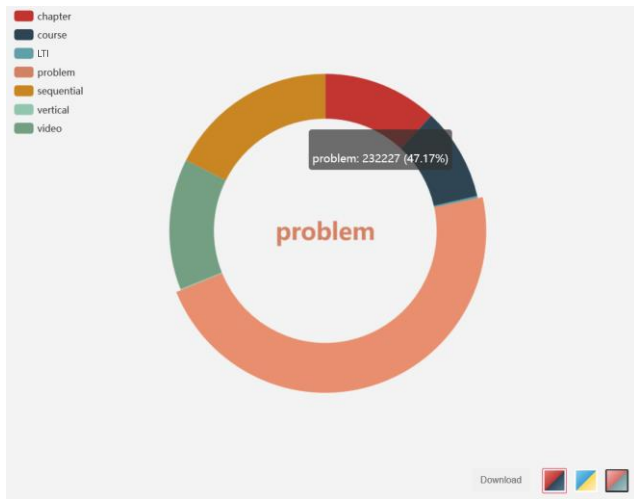


Fig. 1. Different types of learning resource

The distribution of max grade is also demonstrated as Fig.2 shows. The grade values are in the range from 0 to 10. Data of learners with 0 max grade is filtered out since some learners may prefer watching videos to taking tests. 5 and 10 are mostly found as the max grade. There are obvious gaps between score 10, 5 and 1. The performance difference between learners is significant.

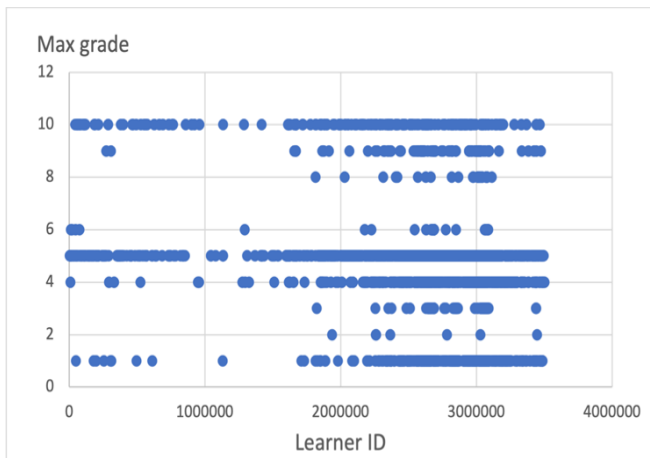


Fig. 2. Max grade of the learners

Learning event number means the times of the learning behaviors. It measures the devotion of learners. Learning event covers various behaviors like slide pages that learners read, marker, and annotate. They have positive effects on the learning performance directly. [9] Learning behavior has relationship with learning performance [10].

The event numbers of different learning resource are demonstrated as Fig.3 shows. Some resource is so difficult that learners must make more efforts to it. Therefore, learners repeat clicking the resource and it has more event numbers. The first few parts of the course have much more event numbers than others. Many learners drop out soon after the start of the course. How to keep the learners' motivation and

make the learners continue studying as they do at the start of the course?

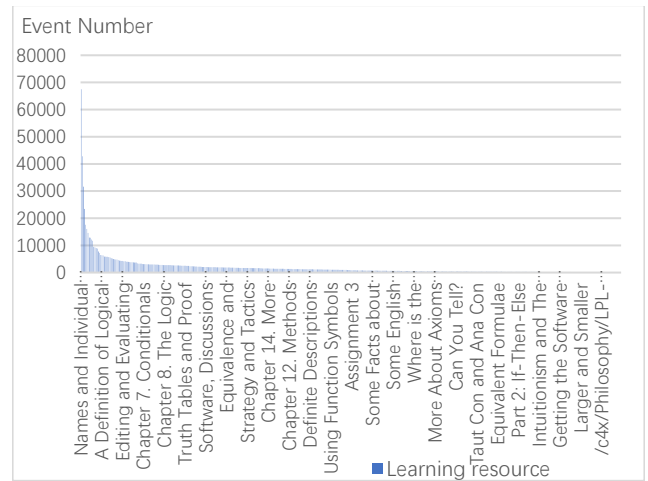


Fig. 3. Event number of different learning resource

IV. A GOOD PERFORMANCE LEARNER CASE ANALYSIS

Learning event data reflects more about the interactions between learners and the platform. To find out what causes the performance difference between learners, event distributions according to diverse factors are analyzed including location, learner, learning time, date, semester, month, resource, etc.

According to the learning data, IP locations are found from many countries as Fig.4 shows. Though it has a long tail of 157 countries, only the top 14 countries are listed due to the page limitation. As English native spoken countries, USA and Canada are mostly found according to the IP locations. Languages and locations are still significant limitations of online learning surroundings. They influence the learners' self-efficacy greatly.

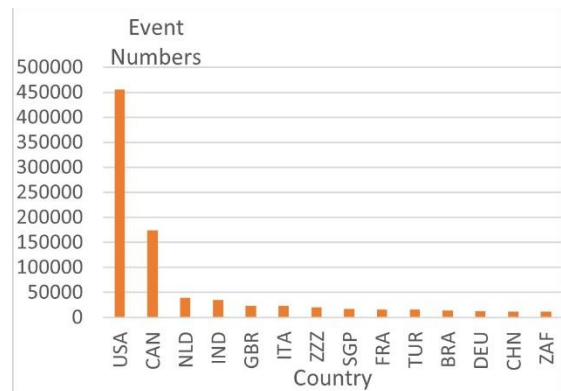


Fig. 4. Distribution of learner countries

Learning event time distribution over different seasons is also listed as Fig.5 shows. Spring and fall semesters have more events than summer and winter semesters. The minimum of

event number is found in winter. Winter and summer semesters have more holidays. More learning events happened in working semesters. Fixed learning time schedule can be followed more easily in working days.

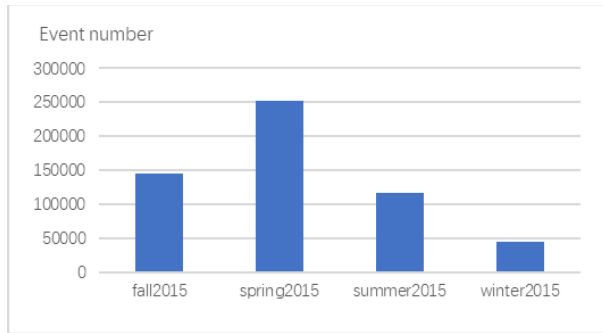


Fig. 5. Event number per semester

Fig.6 is about the learning event numbers in different months. The data of the course in 2015 was adopted in this paper. Holidays were mainly in January and February, so there were hardly any learning events. The course started in March, July, or September according to different semesters. Many learning behaviors occurred in the start months of different semesters. However, event numbers decreased with time after the starts. If the learners don't understand the previous knowledge which are prerequisites to the following learning content, they will be frustrated and lose their patience to continue. The learners will be more likely to drop. [11] Therefore, the event numbers decrease in following months.

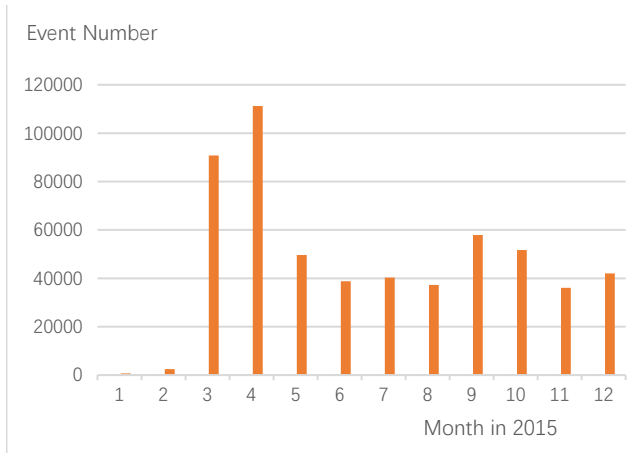


Fig. 6. Event number per month

Fig.7 shows the distribution of learning event numbers over the learners. It is heavily long tailed. There are also great differences between event numbers of learners. While good performance learners finish all the learning resource and achieve high scores, a lot more learners drop out eventually. A series of outstanding examples of excellence in science teaching are presented and the rule of Learning from Excellence is proposed [12]. The data of the good performance learners deserves further analysis.



Fig. 7. Distribution of learning event of different learners

V. A GOOD PERFORMANCE LEARNER SAMPLE ANALYSIS

How can the good performance learners go through the course and persistent in learning? Better performance helps to go through the frustration feeling for better motivation, so that the learner can go on to the end of the course. Motivation is an important aspect of learning because it affects the choices visitors make regarding what to attend to, the amount of effort they devote to learning, and the extent to which they enjoy the experience.[13] Transformation of the learner's identity from a poor learner to a good learner offers an empowering and optimistic way of dealing with difficulties.[14] To analyze the factors for good learning performance, related learners are picked out according to the learning score for further analysis. Here average learning score is adopted as the metric. Learners scored 100 on average are labeled with good performance.

A good performance learner is selected first as a sample for detailed analysis. This learner has a big volume of learning behavior data. It helps understanding the causes of good performance better.

The event numbers of different types of learning resource are listed as Fig.8. The summary parts are the second mostly visited. Learners prefer to achieve more by learning the summary parts. Videos were not clicked for so many times as expected. It shows a good performance learner may go through a video without pausing or repeating. And more interactive learning events about other resource occur. Effective video watching and active interaction like further discussion or revision both contribute to better performance.

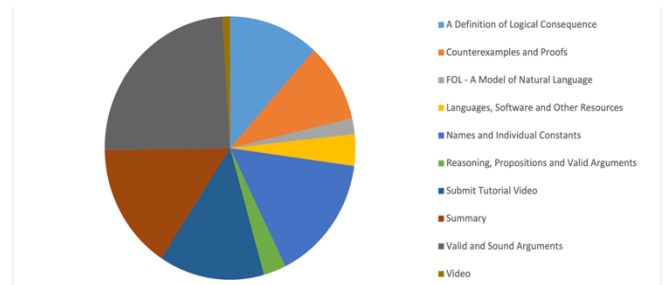


Fig. 8. Event number among different types of learning resource

Fig.9 lists the distribution of event numbers over different learning resource. Most resource was clicked just for one time. Only a little of them was learned repeatedly. It may owe to the learning difficulty. Grasping the video content by viewing it for only one time shows that the good performance learner studies effectively.

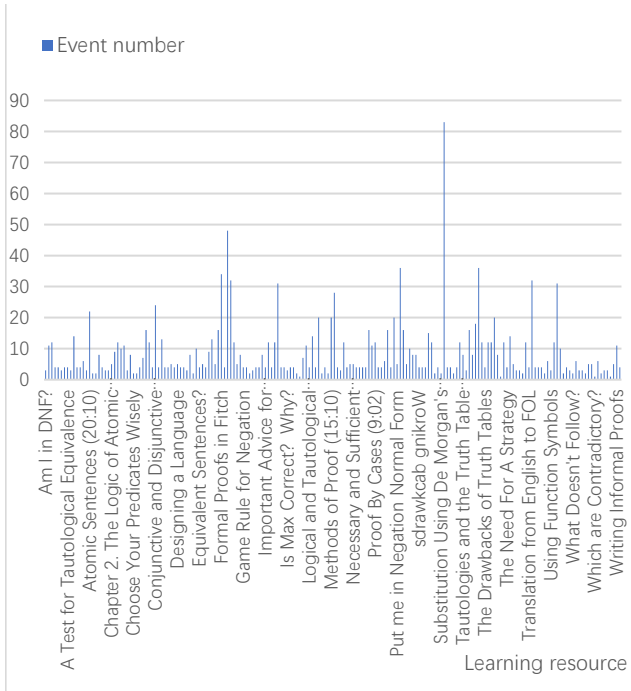


Fig. 9. Event number over different learning resource by a good performance learner sample

Fig.10 shows the distribution of learning time. The horizontal value shows the hour value of the time point. The event number keeps increasing after 16:00 and reaches its peak at 21:00. Then the event number decreases to almost zero at 8:00 in the morning. 20:00 and 22:00 are the turning points. It shows the good performance learners have a relatively fixed learning time schedule. A few hours are kept for learning this course in fixed time span every day. Between 20:00 and 22:00 there is less disturbing. It helps for better concentration on learning.

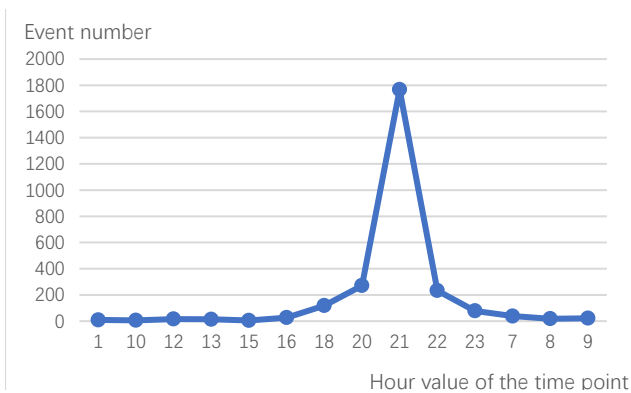


Fig. 10. Event number in different learning event hours of a good performance learner sample

Fixed time scheduling for learning the course daily is suggested for effective work especially in evening time.

VI. TEST EXPERIMENTS ON 4 GOOD PERFORMANCE LEARNERS

To test the analysis result of the good performance learner sample, learning data of another 4 good performance learners is studied in comparison. Their total event numbers are 1121, 2622, 195 and 3942 separately. The sum value of them is 7880.

Compared with the good performance learner sample's data in Fig.9, the learning event distribution of these 4 learners is similar. It is demonstrated as Fig.11 shows. Generally, the learning events also distribute evenly among the resource except for some learning resource that are hard to grasp.

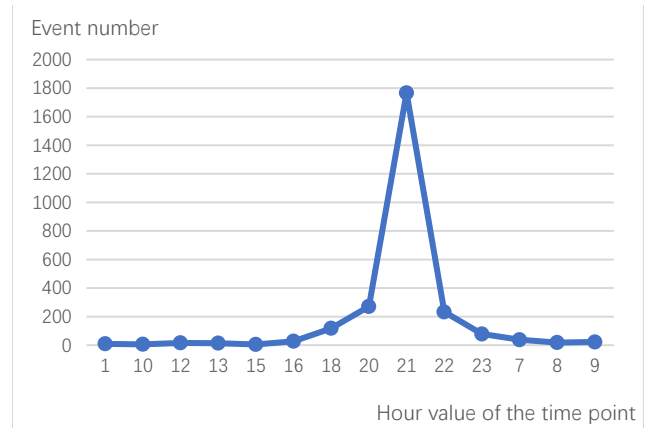


Fig. 11. Event number of different learning resource about the 4 good performance learners

The event time distribution at different learning time point is also tested as Fig.12 demonstrates. The peak time is still at 21:00. Evening time is also popular for learning online according to the good performance learner data. It is verified that fixed time schedule helps for effective learning and contributes to good performance.

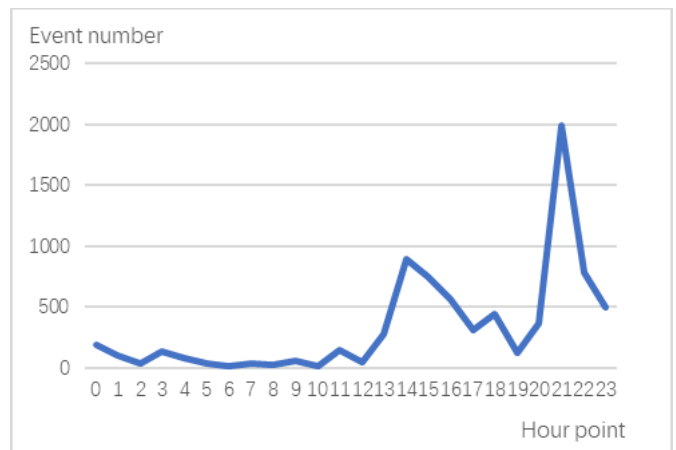


Fig.12 . Event number in different hours of the 4 good performance learners

VII. CONCLUSION

The correlations between learning events and various factors are demonstrated in this paper. English native speaking learners conduct more learning events than those from other countries. Language support enhances the learning environment. In suitable language environment, learners may achieve better performance. Good performance learners usually go through the whole course and work a lot in both video and other learning parts. At the same time, they prefer to learn at fixed time in a day especially in evening time. Learners usually finish video learning effectively at a time. Much time is spared for participation in other parts like discussion, problem answering, summary, etc. That helps for achieving better performance. Contrastive analysis about poor performance learners deserves further research.

ACKNOWLEDGMENT

The course learning data of this paper is provided by the Center for Advanced Research through Online Learning (<https://iriss.stanford.edu>). This paper is supported by the first-class undergraduate profession building of Shanghai under Grant No. A30NH221903-0313.

REFERENCES

- [1] Longhurst, Georga J. , et al. "Strength, Weakness, Opportunity, Threat (SWOT) Analysis of the Adaptations to Anatomical Education in the United Kingdom and Republic of Ireland in Response to the Covid - 19 Pandemic." *Anatomical Sciences Education* 13.3(2020).
- [2] Eom, Sean B. , H. Josephwen , and N. Ashill . "The Determinants of Students' Perceived Learning Outcomes and Satisfaction in University Online Education: An Empirical Investigation*." *Decision Sciences Journal of Innovative Education* 4.2(2010):215-235.
- [3] Kamil Akhuseyinoglu and Peter Brusilovsky, "Data-Driven Modeling of Learners' Individual Differences for Predicting Engagement and Success in Online Learning," *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*. Association for Computing Machinery, New York, USA, 2021, pp.201–212.
- [4] Baber, Hasnan . "Determinants of students' perceived learning outcome and satisfaction in online learning during the pandemic of COVID19." *Asian Online Journal Publishing Group*. 244 Fifth Avenue Suite D42, New York, NY 10001. Fax: 212-591-6094; e-mail: info@asianonlinejournals.com; Web site: <http://www.asianonlinejournals.com> 3(2020).
- [5] U". Cakirođlu. "Analyzing the effect of learning styles and study habits of distance learners on learning performances: A case of an introductory programming course." *International Review of Research in Open & Distance Learning* 15.4(2014):161-185.
- [6] Ashwin, P. . "2005-2010 Centres of Excellence in Teaching and Learning programme." *HEFCE*(2008).
- [7] Pereira, Rsg , et al. "Content analysis as a tool for verifying the learning." (2014).
- [8] Mcklin, T. , et al. "Cognitive presence in web-based learning: A content analysis of student's online discussion." *cognitive processes* 15.1(2002):8.
- [9] Masanori, et al. "Learning analytics of the relationships among self-regulated learning, learning behaviors, and learning performance. " *Research and Practice in Technology Enhanced Learning* (2017).
- [10] Settha Kuama and Usa Intharaksa , "Is Online Learning Suitable for All English Language Students?," *Publications of the Astronomical Society of Australia PASAA*, Vol. 52, July - December 2016, pp.53-82.
- [11] Park, J. H. . "Factors Related to Learner Dropout in Online Learning." *Online Submission* (2007):8.
- [12] John, et al. "Learning from excellence: Some elementary exemplars." *Journal of Elementary Science Education* (1993).
- [13] Packer, J. M. . "Motivational factors and the experience of learning in educational leisure settings." *museum learning* (2002).
- [14] Eliana, H. . "Learning difficulty and learner identity: a symbiotic relationship." *Elt Journal* 1, pp. 33-41(2009).