Design Considerations for Self Paced Interactive Notes on Video Lectures - A Learner's Perspective and Enhancements of Learning Outcome

Suman Deb[1(✉)], Anindya Pal[1], and Paritosh Bhattacharya[2]

[1] Department of CSE, National Institute of Technology Agartala, Jirania, India
`sumandeb.cse@nita.ac.in, anindya2674@gmail.com`
[2] Department of Mathematics, National Institute of Technology Agartala,
Jirania, India
`pari76@rediffmail.com`

Abstract. Video lectures form a primary part of MOOC instruction delivery design. They serve as gateways to draw students into the course. In going over these videos accumulating knowledge, there is a high occurrence of cases [1] where the learner forgets about some of the concepts taught and focus more on what is the minimum amount knowledge needed to carry forward to attempt the quizzes and pass. This is a step backward when we are concerned with giving the learner a learning outcome that seems to bridge the gap between what he knew before and after the course to completion. To address this issue, we are proposing an interaction model that enables the learner to promptly take notes as and when the video is being viewed. The work contains a functional prototype of the application for taking personalized notes from MOOC contents. The work [12] is an integration of several world leading MOOC providers content using application program interface(API) and a customize interface module for searching courses from multiple MOOC providers as COURSEEKA and personalised note taking module as MOOKbook. This paper largely focuses on a learner's perspective towards video based lectures and interaction to find the enhancements in interaction with longer retention of MOOC contents.

Keywords: MOOC note · Self pace learning · Personalized learning
MOOCbook · MOOC Video Interaction · Enhanced learning outcome

1 Introduction

A MOOC is a model of delivering education in varying degrees, massive, open, online, and most importantly, a course [13,14]. Most MOOCs have a structure similar to traditional online higher education counterparts in which students watch lectures online and offline, read material assigned to them, participate in online forums and discussions and complete quizzes and tests on the course material. The

online activities can be supplemented by local meet-ups among students who live near one another (Blended Learning) [3]. The primary form of information delivery in MOOC format is videos. One of the challenges faced by the online learners of today is the need of an interface which enables to take notes from the video lectures [2]. Traditional methods used thus far by the student community are time absorbing and cumbersome in terms of organization. This work is an attempt to address the issue enabling the learner to focus more on the curriculum than on how to compile and access the materials later. As MOOC courses being accessed through out world, beyond any geographical region. Inherently it triggers another level of interaction and understanding difficulty due to cultural, linguistic variation. In addition to that human learning variation takes a great role in graceful MOOC acceptance, learning pleasure and learning outcome.

2   Problem Statement

There is a significant concern over what the learners end up learning as compared to what the MOOC instruction designer intended them to do [4,7]. Many just fall into the trap of knowing just enough to pass the quizzes and course assessments, this neglecting any other concepts that learner may have eventually come across but forgotten about it [8,18]. For the learners who seem to acknowledge this issue on their own, they tend to view the videos again and again until they feel that they have substantial command over the topic being taught in these videos [6,9]. Now, while this may be a good practice, this takes an awful amount of time. Also, watching multiple video lectures on a specific topic may overlap various contents as well tend to forget [15] the previously viewed contents [10,16]. Instead, if there was an interface that lets the learner decide on taking essential parts of the video in a form which can enable them to revise the concepts later and on-demand, it would make sense. This work designs an integrated *MOOC takers note book* that makes an integration of various course providers content on a personalized note interface [11]. This enables cross reference, transcript copy, still frame capture and personalize text note. Taking notes are a manifestation of that conscious effort of peoples natural tendency to forget things with time [19]. Lecture or handouts given in class by the instructor are all the same but people seem to remember more when they are actively taking a record of what is happening, on their own. But there is a flipside to the scenario in digital note taking. People are more reluctant to take notes verbatim, with every word on the document [5]. The trade off between digital and conventional notes are discussed in the experiments presented in [17]. But despite these findings, modern day challenges make to utilize ones time in the best possible way.

3   MOOCbook a Novel Model

Since videos represent the most significant part of MOOCs, it is a mandate that the note taking process will revolve around them. The length of the videos varies from provider to provider, typically ranging from 2–4 min (micro-lectures) to a

maximum of 15 min. As a video progresses, there are certain checkpoints that an instructor breaks a topic into and these checkpoints serve as the keynotes for the topic at hand. For example, a video about supervised learning of machine learning would typically discuss about the common examples in which it is used, then explain the algorithm employed, plot the points representing the features and interpret it, differentiate with other machine learning algorithms and finally conclude the scenarios and advantages where the algorithm applies.
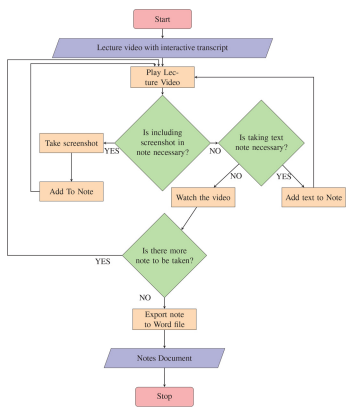
These checkpoints, although important to the MOOC taker at that instant, seem to fade away when the next video starts. The MOOC taker is reluctant on obligating to memory rather tend only to remember those parts which are needed to pass the quizzes. To address this issue, we propose a novel model whereby the MOOC taker can take notes on the fly when they are taking the course through watching videos. For the MOOC taker, the parts of the course which they intend to take note, it happens to be certain points in the video. It is assumed that the video is accompanied by an interactive transcript that scrolls and highlights what the instructor is saying at that moment of the video. During the video, there may happen to be equations, diagrams, graphs and example scenarios that explains the topic from various perspectives. To take the corresponding notes by hand, it would take stopping the video, taking the conventional note book up and writing or drawing whats on the video screen at that instant. This would take up the valuable time that the MOOC taker has invested already. The proposed on the go note taking, while the MOOC taker watches the video is a meta description extraction using a client side scripting on the browser that the learner is currently using to access the materials. The parts of the lecture which catches the attention of the learner are simultaneously displayed in the transcript. A recurrence script extracts transcript with the screen and add the portions to the notebook on events initiated by the user. The learner can save a considerable amount of time which they would otherwise be using for taking the notes conventionally. The user can view the updated note in the browser itself so that it gives a better perspective of what has been learnt (Fig. 1).

Fig. 1. MOOKBook work flow

4 Architectural Design

4.1 Design of COURSESEEKA Module

As a starting point in achieving the goals set forth by, an online interface has been developed where the learners first objective i.e. the need to identify suitable courses that may address his current learning objective, from an array of courses enlisted in various course providers, namely coursera, udacity and udemy. edx
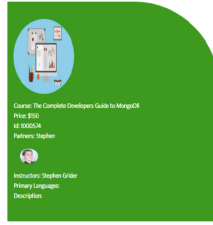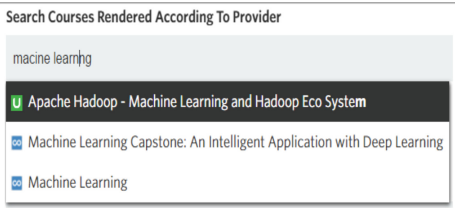
Fig. 2. Fuzzy closeness approximation algorithm in action for filtering a search from multiple MOOC providers simultaneously

Fig. 3. A retrieved course from (a) Udemy (b) Coursera

had also been approached for their API on two occasions but both the requests got rejected. These course providers are fairly popular and have gained trust among the learning masses as the MOOC movement took place. Also, these have well defined APIs which enlist course related information that can be obtained easily. The COURSEEKA interface is based on the architecture as described by Fig. 5. The interface aims to find courses available from three course providers, namely courser.org, udacity.com and udemy.com and combine their results into a single web page where a user can query a course specific search term according to his learning objective, and the courses will then be filtered accordingly (Figs. 2, 3, 4 and 5).
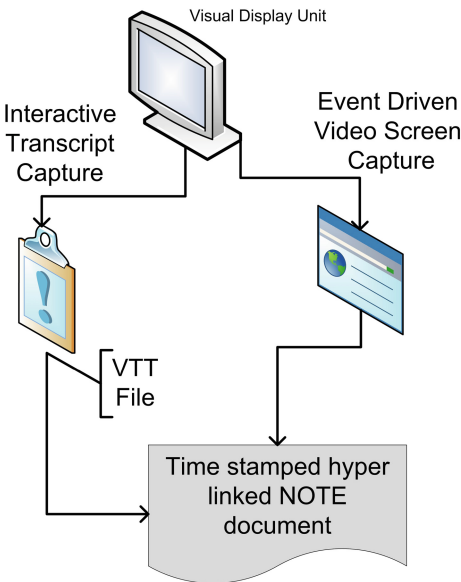

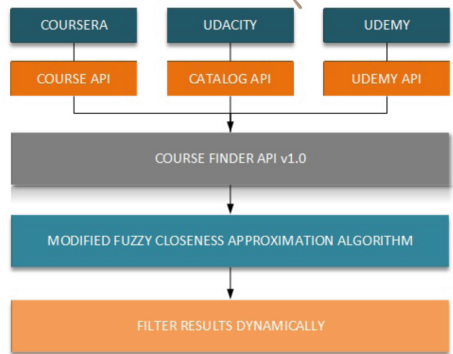**Fig. 4.** MOOKBook multi modal note generation

Fig. 5. COURSEEKA API stack

4.2 Modified Fuzzy Closeness Approximation Algorithm

Existing interfaces on course search is based on matching the keywords wholly. While this may seem as a very nave way to get courses recommended to a learner based on his search term, our web application has a learner centric approach to getting the search results that will suit someone who is willing to manage his online MOOC curriculum in a very specific way. His search results are constrained to be from one of the major MOOC providers (as has been told already) existing today. Moreover, the search algorithm is based on a modified fuzzy string closeness approximation algorithm which is clever enough to infer what the MOOC learner is specifically searching for even if he is halfway through or even less than what he intends to type.

5 Implementation

5.1 Prototype Specifications

The prototype is a web application that hosts a video with interactive transcript and has control buttons to preview and append notes. The interface aims to capture portions of the text of the video content i.e. the transcript along with screen captures, preview them and append to the notebook inside the webpage itself. Finally, the user has the option to download the notebook thus formed. All of this happens using client side scripting, which is relevant since time is of the essence when the user is taking the note as the video plays. This eliminates the load off the servers hosting massive amounts of data in the MOOC servers.

5.2 Prototype Demonstration

An initial working prototype has been implemented which uses the three APIs combined and lists all the courses relevant to a learners interest as they types in a search query. The search results are then displayed centrally using a Fuzzy String Closeness Approximation. As an example of working demo, the video course cited is one of the those featured in the first week of the Machine Learning course by Professor Andrew Ng of Stanford university, hosted by coursera.org. The instructor goes about explaining Un-supervised learning in the course. Figure 6 The distinguishable parts of the video are listed as under: 1. Difference between unsupervised learning and supervised learning (two graphs). 2. Applications of Supervised Learning (images depicting them). 3. Tackling a problem (cocktail party problem) using unsupervised learning (image de-picting the scenario). 4. Cocktail party problem algorithm (code in python) 5. A quiz with options to choose from. These distinguishable parts are of concern to the learner when compiling a digital note about the video. The MOOCbook interface is equipped to take snapshots of these parts and scrape the transcripts of the relevant portions as and when the learner deems it necessary. Figure 6 shows a screen of the video captured for preview. The snapshot is taken using the videos and videos interactive transcript JS libraries in tandem. If the preview is deemed good for
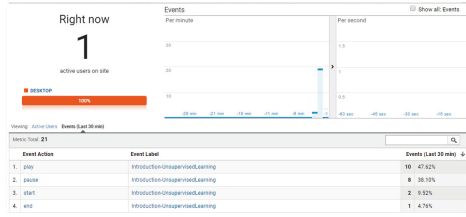
Fig. 6. MOOKBook GUI and interactions
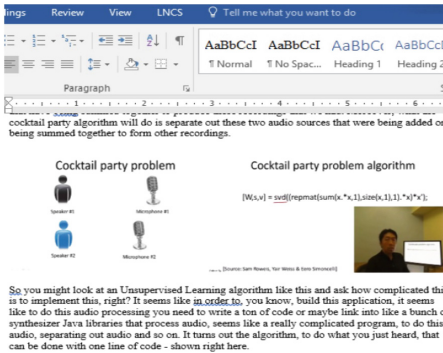

Fig. 7. Analytic dashboard

Fig. 8. MOOCBook final note in MS word

Fig. 9. Example clickstream data collected from google analytics

Event Action	Total Events	Unique Events	Event Value	Avg. Value
percent played	18	1	900	50
play	11	1	2,456	223.27
seek end	7	1	3187	455.29
seek start	7	1	2032	290.29
pause	4	1	62	15.5
start	1	1	0	0
end	2	1	0	0
volume change	1	1	0	0
enter fullscreen	1	1	0	0
exit fullscreen	1	1	0	0
	53	8	8637	162.96

adding to the note, the user then proceeds accordingly. To capture the lecture discussions relevant to the note being compiled, we have made use of the VTT file available with the video in courser. The VTT file has timestamps along with text content, which is scraped using suitable javascript code, and added to the note. Thus, the cocktail party problem algorithm now has a proposed problem, a solution with code and relevant transcripts, all in one note, viewable in the browser itself where the video is still playing. The note thus far compiled, is now available for download to the client machine using the jquery word export plugin made using JS. The final note file is a MS Word document (Figs. 7, 8 and 9).

6   Synthesis of Experiments and Result

The system developed for taking notes from MOOCs, namely MOOCbook is taken up for testing effectiveness. Pretests were concluded before the actual experiment to establish clear reference point of comparison between treatment group and control group. To investigate whether the proposed system effectively generates a learning outcome that lasts even after the video completes, post tests were conducted between the two groups. The subject matter that is portrayed in

the two videos which are featured in the system developed is an introduction to the two major varieties of Machine Learning algorithms. Both the treatment and the control groups have a basic knowledge of what Machine Learning is about.

6.1   Evaluation Criterion

The current MOOC interfaces available on the Internet featured on MOOC platforms like coursera, udacity etc. are designed to deliver content over multiple media formats. The primary format, namely videos are designed to be accompanied by in-video quizzes that assess the learners comprehension with the help of in-video quizzes as well as separate assessments module. But certain parts of the video are overlooked by the learner because he may be impulsively following the video to complete the quizzes and the assessments. For this purpose, it happens that the learner may have peeked into the questions beforehand and accordingly is inclined to get the answers from the video. So, he is skimming portions of the video in order to find the answers and thus is not open to effective learning. The questions to understand how the system enhances the learning outcome of a learner have been identified as under:

– *Question 1:* How much time is spent on viewing the video, including activation of play and pause buttons on the video player?
– *Question 2:* Whether skimming the video helps in understanding the content?
– *Question 3:* Did the users feel the need to seek to certain parts of the video to find answers to questions which are known beforehand the experiment?
– *Question 4:* Does a digital note assistant help in reviewing and recalling portions of the content in a way that saves time and thus increase the effective learning of the MOOC taker?
– *Question 5:* Did the users who were provided with the MOOCbook module actually refer the downloaded note?

6.2   Methodology

The participants of this experiment are 6th Semester Under Graduate Engineering students. There are 84 students in total, divided into two groups, one being a control group and the other being the treatment group. They are shown two videos each on the system developed. The control group gets to see only the videos, while the treatment group sees the MOOCbook interface at play, which enables them to take notes if necessary. Each of the participants are allotted 40 min for viewing the videos. The combined length of the videos is $(12.29 + 14.13) = 26.42$ min. Throughout the duration of the video featured in MOOCbook, all activities of the user are recorded with the help of Google analytics that serve as a gateway to learn key insights into how the users interact with the video player while seeing the videos. The data collected through Google analytics is downloadable and hence form our dataset of study. The data downloaded from Google analytics is in the form of csv files which are obtained individually from all the 84 users of the experiment. The effectiveness of the

MOOCbook interface was tested using independent-samples t-test. It is aimed to compare means between two unrelated groups on the same continuous variable. In this case, it has been used to understand whether the learning outcome undergraduate engineering student is increased on the application of MOOCbook. Thus the independent variable here is "User of MOOCbook or not "(one of the groups being users who had MOOCbook interface at their disposal and the other being who did not use MOOCbook) and the dependent variable is the "learning outcome ".

Assumptions. As requirement of independent t-test, the 6 point compliance of assumptions as detailed under.

1. The dependent variable, namely learning outcome is measurable in a continuous scale.
2. The independent variable i.e. whether MOOCbook user or not, has two possibilities viz. either the user is given the MOOCbook interface or the user is not. Thus there are two categorical, independent groups.
3. There is independence of observations since there is no relationship between the observations in each group or between the groups themselves.
4. There is no significant outliers, meaning there are no values in the dataset that does not follow the usual pattern.
5. The dependent variable which is learning outcome is approximately normally distributed for each group of the independent variable.
6. There is homogeneity of variances.

6.3   Instruments

The various analytical processes aimed at answering the questions identified have been enlisted here. A short demonstration was performed which walked through the MOOCbook interface to the participants before the experiment so that they are familiar with the system. A questionnaire aimed at measuring MOOC awareness among the participants serves as a pretest before the experiment, and two post tests comprising data analysis of clickstream events generated during experiment and a quiz is aimed at testing effectiveness of the MOOCbook interface.

Pretest. The pretest was carried out before the participants were given access to the system. The two groups were surveyed about their MOOC awareness. A questionnaire specific to MOOC awareness was used in this regard.

Post Intervention Tests

1. **Clickstream data analysis** - To address how behavior of participants differ on the provision of the MOOCbook interface in terms of interaction with the video (questions 1–3 of **Evaluation Criteria** section), the data generated through clickstream events of the video on the google analytics server was analyzed.

2. **Learning outcome** - To answer the questions 4 and 5 enlisted in **Evaluation Criteria** section, a quiz was conducted with the participants and the results were evaluated.
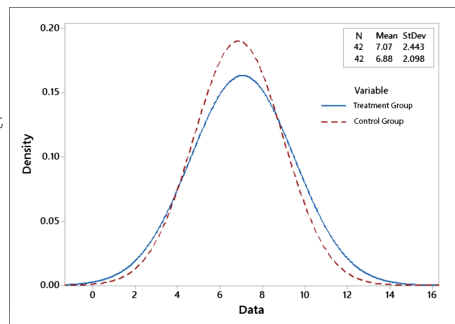
Null Hypotheses

– H1: There is no significant difference in terms of MOOC taking experience between the treatment group and the control group.
– H2: There is no significant difference between the participants in terms of the pattern of clickstream events generated for the videos watched.
– H3: There is no significant difference between treatment and the control group in terms of learning outcome generated from the experiment.

Pretest Results. Data in Fig. 10 shows that the treatment group pre-test mean scores was 7.07 (SD = 2.443) while the control group pre-test mean score was 6.88 (SD = 2.098). To ensure the comparison between two groups, a two tailed t-test was done on the sample for 5% level of significance. The findings are shown in Fig. 11. The mean difference between the treatment and the control group in terms of Test Scores is 0.190. The findings (Fig. 11) lead to the conclusion that there is no significant difference between the treatment and control group prior to the experimental study conducted. Both groups were found to have common ground of knowledge when it comes to MOOCs and thus are ideal for the MOOCbook test scenario. Hence hypothesis H1 failed to be rejected (Fig. 12).

Variable	Observations	Minimum	Maximum	Mean	Std. Deviation
Treatment Group	42	3	10	7.07	2.443
Control Group	42	3	10	6.88	2.098

Fig. 10. Pre-test score results of treatment and control group

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Test Scores	Equal variances assumed	1.533	0.219	0.383	82	0.702	0.190	0.497	-0.798	1.179
	Equal variances not assumed			0.383	80.162	0.702	0.190	0.497	-0.798	1.179

Fig. 11. Independent samples test as pretest

Fig. 12. Normal distribution for pretest scores

Post Test Results Clickstream Data Analysis. Data in Fig. 13 shows the summary of clickstream data obtained from 84 participants. The control group generated analytics data from only the video player interactions like play, pause, fullscreen etc. while the treatment group was capable of generating note-taking events like Add Text To Note, Add Image To Note etc. in addition to what control

group users were allowed. For the purpose of analysis, only the clickstream data with respect to video player interactions is taken up. The note-taking interactions will not be taken in the post test analysis. The normal distribution graph of the post-test-1 scores for the two groups is shown in Fig. 15. To analyze the hypothesis H2, a two tailed t-test was done on the sample for 5 % level of significance. The mean difference between the treatment and the control group in terms of the number of events registered while watching the videos is -32.667. The findings of the Independent Samples Test is depicted in Tabular data Fig. 14. The above findings lead to the conclusion that there is a significant difference between the treatment and control group post the experimental study conducted. Both groups were found to have interacted in a very different way when it came to viewing the videos. The number of clickstream events were far higher for the control group without the notes system than the treatment group with notes enabled. This leads to the conclusion that hypothesis H2 is false and does not hold.

Variable	Observations	Minimum	Maximum	Mean	Std. Deviation
Treatment Group	42	18	63	40.24	14.043
Control Group	42	38	129	72.90	21.354

Fig. 13. Post-test clickstream results of treatment and control group

Fig. 15. Normal distribution for post test 1 scores

		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference
									Lower	Upper
Events	Equal variances assumed	4.100	0.046	-8.283	82	0.000	-32.667	3.944	-40.512	-24.821
	Equal variances not assumed			-8.283	70.876	0.000	-32.667	3.944	-40.530	-24.803

Fig. 14. Independent samples test as post-test 1

Learning Outcome. The Post Test 2 is a questionnaire that aims to find the learning outcome of the participants. The questions contained here are set from the content of the two videos that are hosted in the MOOCbook system. The control group once again is devoid of the functionality of taking notes whereas the treatment group is notes module enabled. The results obtained as shown in Fig. 16 will be directly connected with how much of the lessons depicted within the videos are comprehended by the users. Thus the direct measure of how much a knowledge a learner can retain will be obtained. To analyze the hypothesis H2, a two tailed t-test was done on the sample for 5% level of significance. The mean difference between the treatment and the control group in terms of the Qscores (scores obtained by the participants on the questionnaire) is -2.333. The findings of the Independent Samples Test is depicted in Fig. 17. The findings lead to the conclusion that there is a significant difference between the treatment and control

Variable	Observations	Minimum	Maximum	Mean	Std. Deviation
Treatment Group	42	-5	8	40.24	14.043
Control Group	42	-2	10	72.90	21.354

Fig. 16. Post-test 2 results of treatment and control group

		Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
QScores	Equal variances assumed	4.915	0.029	-2.951	82	0.004	-2.33333	0.79070	-3.90630	-0.76037
	Equal variances not assumed			-2.951	75.697	0.004	-2.33333	0.79070	-3.90826	-0.75841

Fig. 17. Independent samples test as post-test 2

group post the experimental study conducted. Both groups were found to have had a very different learning outcome in terms of understanding the contents depicted in the videos. The number of correct answers for the quiz questions were far higher for the treatment group with the notes system enabled than the control group with notes disabled. This leads to the conclusion that hypothesis H3 is false and does not hold. Thus the notes module plays a significant part in terms of making the lessons more content aware to the learners. They are able to differentiate key points told by the lecturer and form memory mappings of lesson checkpoints which later help them to retrieve the same, i.e. recall lesson key points.

7   Conclusion

This work is an attempt to address the issues enabling the learner to focus more on the curriculum than on how to compile and access the materials later. A novel model MOOCbook was presented and a working prototype has been demonstrated for this purpose. The results obtained have provided us with some insights to get into what people are looking for in terms of enhancing their learning outcome. One of the major finding was a need of self paced MOOC note. The empirical experiments conducted and anecdotal response have shown significant improvement in engagement to accomplish MOOC course as well enhancement in learning outcome. All the work has been done from a learner's perspective. The inclusion of this tool along with MOOC provider's platforms will pave the way for enhanced digital learning in the future.

References

1. Barba, N., Gluhanich, L.: 4 Pros and 4 Cons of MOOCs: Whether to take study from classroom to online (2014). https://www.geteverwise.com/career-success/4-pros-and-4-cons-of-moocs-whether-to-take-study-from-classroom-to-online/. Retrieved 19 June 2017
2. Giannakos, M.N., Jaccheri, L., Krogstie, J.: Looking at MOOCs rapid growth through the lens of video-based learning research. IJET **9**(1), 35–38 (2014)
3. Davis, K., et al.: The theory of multiple intelligence. In: The Cambridge Handbook of Intelligence, pp. 485–503 (2011)
4. Fleming, N., Baume, D.: Learning Styles Again: VARKing up the right tree!. Educational Developments 7.4 (2006)
5. Barrington, E.: Teaching to student diversity in higher education: How multiple intelligence theory can help. Teaching High. Educ. **9**(4), 421–434 (2004)
6. Redding, S.: Through the Student's Eyes: A Perspective on Personalized Learning and Practice Guide for Teachers. Temple University, Center on Innovations in Learning (2013)
7. Kop, R., Fournier, H., Mak, J.S.F.: A pedagogy of abundance or a pedagogy to support human beings? Participant support on massive open online courses. In: The International Review of Research in Open and Distributed Learning, vol. 12(7), pp. 74–93 (2011)
8. Vivian, R., Falkner, K., Falkner, N.: Addressing the challenges of a new digital technologies curriculum: MOOCs as a scalable solution for teacher professional development. Res. Learn. Technol. 22 (2014)
9. Norton, A., Sonnemann, J., McGannon, C.: The Online Evolution: When Technology Meets Tradition in Higher Education, vol. 48. Grattan Institute, Australia (2013)
10. Kostolnyov, K., Armanov, J.: Use of adaptive study material in education in E-learning environment. Electron. J. e-Learn. **12**(2), 172–182 (2014)
11. Gynther, K.: Design framework for an adaptive MOOC enhanced by blended learning supplementary training and personalized learning for teacher professional development. Electro. J. e-Learn. **14**(1), 15–30 (2016)
12. Deb, S., et al.: Enhancing personalized learning with interactive note taking on video lectures-an analysis of effective HCI design on learning outcome. In: Communications in Computer and Information Science (CCIS). Springer (2017). ISSN No: 1865–0929(2017)
13. IITBombayX: Iitbombayx.in (2017). https://www.iitbombayx.in/dashboard. Retrieved 27 June 2017
14. NMEICT: Nmeict.iitkgp.ernet.in (2017). http://www.nmeict.iitkgp.ernet.in/audio.php. Retrieved 27 June 2017
15. Rubin, D.C., Wenzel, A.E.: One hundred years of forgetting: A quantitative description of retention. Psychol. Rev. **103**(4), 734 (1996)
16. Murre, J.M., Chessa, A.G., Meeter, M.: A mathematical model of forgetting and amnesia. Front. Psychol. **4**, 76 (2013)
17. Mueller, P.A., Oppenheimer, D.M.: The pen is mightier than the keyboard advantages of longhand over laptop note taking. Psychol. Sci. **25**(6), 1159–1168 (2014)
18. Clow, D.: MOOCs and the funnel of participation. In: Proceedings of the Third International Conference on Learning Analytics and Knowledge. ACM (2013)
19. Ros, C.P., et al.: Social factors that contribute to attrition in MOOCs. In: Proceedings of the first ACM Conference on Learning @ Scale Conference. ACM (2014)