Number 1

1 A design-based approach to support and nurture open educational practices
Shironica P. Karunanayaka and Som Naidu

21 Learning analytics in higher education: an analysis of case studies
Billy Tak Ming Wong

41 Student characteristics and knowledge on ODL concepts at first registration: a case study from OUSSL
K.A.J.M. Kuruppuarachchi and K.O.L.C. Karunanayake

52 Analyzing students online learning behavior in blended courses using Moodle
Rosalina Rebucas Estacio and Rodolfo Callanta Raga Jr

69 Quality assurance in learning material development at OUM
Nazrai Ahmad Zabidi, Tai Kwan Woo, P. Rajesh Kumar, Mansor Fadzil and
Syafiah Hidayatul Syed Husain

82 The online examinations at Universitas Terbuka: an innovation diffusion viewpoint
Sugilar Sugilar

94 The potential and challenges of MOOCs in Pakistan: a perspective of students and faculty
Syed Saad Ahmed, Eisa Khan, Muhammad Faisal and Sara Khan

106 A multi-objective evolutionary algorithm-based soft computing model for educational
data mining: a distance learning experience
Choo Jun Tan, Ting Yee Lim, Chin Wei Bong and Tek Kooi Liew
A design-based approach to support and nurture open educational practices

Shironica P. Karunanayaka
Faculty of Education, The Open University of Sri Lanka, Nawala, Sri Lanka, and
Som Naidu
The University of the South Pacific, Suva, Fiji

Abstract

Purpose – A critical attribute of open educational practices (OEP) is the pursuit of open scholarship which comprises the release of educational resources under an open licence scheme that permits no-cost access, use, reuse, adaptation, retention and redistribution to others. The degree of openness in relation to this attribute will depend on the context and culture of the place and the people in it. When left to chance, the adoption and practice of open scholarship by educators is at best sketchy. For optimum impact, a design-based approach is essential. A central focus of such an approach will need to target educators’ belief systems and practices about their scholarship. Any such work will involve researchers collaborating with practitioners in real-life settings to improve educational practices through iterative analysis, design, development and implementation. The purpose of this paper is to report on how the development and use of such a design-based approach, implemented by the Open University of Sri Lanka, impacted the adoption and uptake of open scholarship among teachers in the Sri Lankan school system in terms of changes in their use of instructional resources, pedagogical thinking and pedagogical practices.

Design/methodology/approach – The study adopted a design-based research (DBR) approach (Reeves, 2006), which involved researchers collaboratively working with practitioners in real-life settings to improve their educational practices along three aspects – instructional resource use, pedagogical perspectives and pedagogical practices. Based on the four stages of the DBR approach – analysis, solution, testing and refinement, and reflection, a professional development intervention programme was designed and implemented to support teachers on the integration of open educational resources (OER) and adoption of OEP in their teaching-learning process. Data collected throughout the process using multiple strategies such as questionnaire surveys, concept mapping, lesson plans, focus group interviews, self-reflections and “stories”, were analyzed using both qualitative and quantitative methods.

Findings – By the end of the intervention, significant changes were observed in teachers’ use of instructional resources, their pedagogical thinking and pedagogical practices. While resource usage has shifted from no or low usage of OER to reuse, revise, remix and creation of OER, the pedagogical thinking and practices of teachers moved from a content-centric and individualized patterns to more constructivist, context centric and collaborative ways. The diffusion of OEP was prominent along two dimensions – enhancements in the individual practices in innovative OER use as well as collaborative practices of sharing of resources, knowledge and good practices.

Practical implications – The systematic and flexible methodology adopted based on the DBR approach via a framework designed as a contextualized, process oriented and self-reflective enquiry has been very useful to support changes in OEP among practitioners over time.

Originality/value – This iterative process allowed the researchers to function as “designers”, while investigating real-life issues in collaboration with the practitioners through reflective enquiry to further refine
innovative practices towards OEP. This provides valuable insights for improved design solutions for future interventions in similar contexts.

**Keywords** Teacher professional development, Open educational resources, Design-based research approach, Open educational practices, Open scholarship

**Paper type** Research paper

## Introduction

The imperative for open educational practices (OEP) has been on the rise for some time now. Foremost, this comprises; open access to educational opportunities; open learning, at the heart of which is time, place and pace flexibility; and open scholarship which means free and open access to educational resources for use, reuse, adaptation, retention and redistribution to others. The concept of openness is not new. Open access to educational opportunities and flexible approaches to learning and teaching have always been a core defining characteristic of open, flexible and distance learning which sees openness and flexibility in relation to time, place and pace of learning as a value principle, much like we see diversity, equity or equality in education and society more broadly. Openness and flexibility are at the heart of what distance educators do and promote. And in the contemporary world, both in the developed and developing contexts, technological infrastructure is a key component of it.

The notion of open scholarship is a more recent phenomenon and an addition to the broader concept of openness. Its value principle is that education in general, and knowledge more specifically, is a public good which should be available to all. A key component of this is the release of educational resources under an open licensing scheme (e.g. Creative Commons Licensing Framework) as open educational resources (OER) for it to be able to be used, reused and shared freely and openly at no cost to either party. The imperatives for adopting this path to open scholarship are enormous. For teachers and learners, both in developed and developed educational contexts, this means access to valuable and adequate educational resources for better teaching and learning. And it includes the opportunity to adopt and adapt these essential educational resources in ways to suit local contexts and its requirements. These affordances are critical and crucial for the achievement of our sustainable development goals towards education for all and promoting equity, and equality of educational opportunity more broadly but especially in developing social and economic contexts.

While the educational imperatives for open scholarship are clear and convincing, its adoption and implementation more widely are still fraught with challenges. Some of these challenges have to do with the economics of open scholarship. Someone, and at some point, must pay for the production of educational resources so that these can be distributed widely and openly among the community, and at no cost to its users. Issues around this kind affordability are still being worked out by authors, developers and publishers of any such content.

The more serious challenge has to do with the adoption and integration of OER in teaching. Foremost this requires an understanding of the concept of open scholarship and what it involves. This includes an understanding of what is an OER, where is it found, and how to find it? How is it distinguishable from any other kind of educational resource? And also, how to develop an OER? What, if anything, is qualitatively different about it? And once found or developed, how best to make use of it in supporting teaching and learning? What are its pedagogical affordances? And how these affordances differ from those that are possible for proprietary educational resources?

A lot of this has to do with developing new perceptions, and perspectives about pedagogy, about what it means to teach and to learn. And this involves, not just developing new and technical skills but it is about shifting mindsets. It requires practitioners to move beyond a focus on access to OER, to a more comprehensive view about creation and integration of OER in order to make a difference in the teaching-learning process. This requires engagement of teachers and learners in open practices
which are participatory, collaborative and innovative, the extent of which will vary, depending on the context and culture of places and people. Far too many attempts at the integration of OER have failed to impress upon teachers and learners the pedagogical affordances of use, reuse and adaptation of OER. Many have failed to make clear to teachers and learners that not all educational resources, no matter how good, are fixed or static entities. Surely, some of this content is declarative and does not change, but others do change and need to be seen differently in different contexts. As such focussing on the learning of content cannot be the end game for learners. A focus on learning to solve real-life challenges and problems should be the focus of all learning. The subject matter content for this can and should be sourced from anywhere. And it helps if these are OER so that they can used, reused and manipulated in ways that are not possible with proprietary material. Such an approach requires a shift in the mindsets of educators and teachers from thinking about teaching to the content to designing rich and relevant learning experiences in which students are engaged in solving real-world challenges and problems with the help of OER.

The design of such learning experiences requires a close working relationship and partnership between researchers and practitioners in the field. In our case, these practitioners are the teachers in the Sri Lankan school system. A core component of this partnership is the joint development of learning experience designs by researchers and practicing teachers, such that both parties can claim a sense of ownership of what is being developed, and it is not something that is being imposed upon them. Such a working relationship is at the heart of a design-based approach. It involves researchers and practitioners in the field working closely to design artefacts that can be implemented and data collected on its effectiveness through iterative analysis, design, development and implementation in situ.

This paper describes a programme of work along these lines that has been carried out in Sri Lanka over the period 2015-2016. This project was led by researchers from the Open University of Sri Lanka (OUSL), with funding from the International Development Research Centre, Canada, and administered by Wawasan Open University, Malaysia and the University of Cape Town, South Africa. The goal of this work has been to study the impacts of the adoption of OER and OEP by school teachers in Sri Lanka. This paper reports on the use of design-based research (DBR) in supporting and nurturing the adoption of OEP by teachers in the Sri Lankan school system in terms of changes in their use of educational resources, their pedagogical perspectives and pedagogical practices.

Review of literature

Opening up education through OEP

"Opening up education" is a concept that has been given a great deal of attention during the recent past, as witnessed by a continuum of “open” concepts and practices in relation to teaching and learning that have been evolving over time. Under a larger agenda of open education which emphasizes that knowledge should be shared for the greater good of the community, there are various dimensions of open concepts such as open learning, open teaching, open access, open scholarship, OER and OEP, supporting this agenda. All these concepts focus on enhancing “openness” in education by removing barriers to learning and addressing student needs, providing opportunities for lifelong learning and flexibility over when, where, how and at what pace to engage in study, improving access to knowledge, enabling efficient and affordable sharing of educational resources and offering alternative ways of teaching and learning (Butcher, 2011; COL, 2000; Naidu, 2016; Wiley and Green, 2012).

Openness in teaching and learning practices is based on the value principle that knowledge should be shared – freely, openly and equitably. It signifies that knowledge, as a common good, should be accessible and usable by all human beings. Opening up access to educational resources, and opening up practices in the use of such resources are
two focal areas of openness in education. The release of teaching, learning and research materials under an open licensing framework (Creative Commons, n.d.) as OER has increased no-cost access to and sharing of knowledge in an efficient and affordable manner (UNESCO-COL, 2012; Wiley and Green, 2012). This has enhanced emergence and adoption of various scholarly practices of openness, promoting open scholarship, which is a critical attribute of OEP.

The “spectrum of rights” provided by the open licensing framework to users giving access to resources through “most open” to “least open” licence types, allows an opportunity to adopt varying degrees of openness in the use of educational resources via the 5R framework of OER – retain, reuse, revise, remix and redistribution (Wiley, 2014). The ability to adapt and use educational resources in this manner, which they were unable to do with the proprietary materials, has empowered educators to become more creative and innovative in their educational practices. In other words, the OER movement has stimulated innovations in teaching and learning practices. This shift of focus from access to resources, to innovative practices in the use of resources promotes the concept of OEP (Ehlers, 2011; Open Educational Quality Initiative Report, 2009).

The various definitions of OEP fundamentally focus on how “openness” can be practiced in teaching and learning. For instance, OEP are explained as, a set of activities and support around the creation, use and re-purposing of OER (Conole, 2010); a combination of open resource use and open learning architectures to transform learning (Camilleri and Ehlers, 2011); use and production of OER in such a manner to improve the quality of education and innovate educational environments (Ehlers and Conole, 2010); and practices that support the (re)use and production of OER through institutional policies, promote innovative pedagogical models, and respect and empower learners as co-producers on their lifelong learning path (Ehlers, 2011). The emphasis on practices instead of resources and on enhancing innovations and quality in education are common features stressed in all these definitions.

These practices are further described in relation to the use of educational content through “open pedagogies” and “open technologies” (Beetham et al., 2012). While the adoption of pedagogical models based on more constructivist and participatory approaches will allow educators to become more open in their teaching practices, increased availability of free and open technological applications and services let educators select appropriate technology to support innovative practices. OEP will thus require engagement of teachers and learners in innovative practices which are participatory and collaborative, harnessing the potential of both technology and pedagogy (Cape Town Declaration, 2008). However, such changes in practices would essentially require changes in practitioners’ belief systems and mindsets. Hence, promoting adoption of open scholarship and OEP becomes very challenging.

Facing challenges in implementing OEP

The degree of openness in educational practices, and hence open scholarship will vary, depending on the context and culture of places and people. While the degree of openness in actual content will differ based on their licence type, the degree of openness in practices will change depending on the pedagogical models practitioners adopt in the use of the educational resources. In order to make an impact, implementation of OEP would essentially require practitioners to change from a simple focus on access to a broader view about integration of OER in the teaching-learning process and creation of OER, which is a challenging process (Ehlers and Conole, 2010). Even though adoption of OER enhances access to resources, OER by themselves would not have any impact unless those are used in particular ways for a change in practices to occur (Smith, 2016). This will require practitioners using innovative methods in the integration of OER. Research studies conducted on OER integration in teaching and learning reveal that adoption of OEP demands a culture of sharing and use of novel pedagogical models (Beetham et al., 2012;
The realization of successful OEP will thus depend on having opportunities for pedagogical innovation, yet innovative thinking and practices may occur if and when more flexibility and choices are available for practitioners (Educause, 2010; Weller, 2014). Five principles of openness necessary for implementing OEP have been identified, comprising open tools and processes that will promote: collaboration and sharing of information; connected communication about learning and teaching; collectivity to grow knowledge and resources; critique for the promotion of scholarship; and serendipitous innovation (Conole, 2013). Digital technologies will play a key role in achieving such requirements effectively and efficiently and hence digital competencies also becomes a crucial requirement in supporting OEP (Beetham et al., 2012; European Commission, 2013).

Various initiatives on OER integrations have presented frameworks for implementing OEP in different contexts. For instance, the “OPAL framework” (Andrade et al., 2011), provides strategies to consider when designing, developing, implementing and evaluating OER initiatives. Capturing the link between resources and practices, OEP is explained in two dimensions in this framework: openness in resource usage and creation versus openness in pedagogical models, suggesting different degrees of openness in both aspects from low to high levels (Ehlers, 2011). Similarly, the 7C learning design framework (Conole, 2014) presents seven elements: conceptualize; capture; communicate; consider; combine and consolidate, integrating the trend for openness by using OER and collaborative practices. A model of “open pedagogy” discusses eight interconnected attributes – participatory technologies; people, openness, trust; innovation and creativity; sharing ideas and resources, connected community; learner generated; reflective practice and peer review (Hegarty, 2015). Such models exemplify how OEP can be facilitated by creating structured enabling environments.

It is apparent that adoption of OER and OEP by practitioners can be truly effective only if it reflects a change in their thinking and actions. Educational change is a complex process, and to deal with such complexity, the best way is not to control change but to guide it, by individuals taking action as “change agents” (Fullan, 1993). Enacting change towards OEP is best achievable through a strategic approach via systematic design of appropriate learning experiences (Naidu and Karunanayaka, 2014). OER integration endeavours in the Sri Lankan context, have revealed significant influences in supporting changes in thinking and practices of educators towards OEP through design and implementation of effective, efficient and engaging learning experiences based on innovative pedagogical models (Karunanayaka and Naidu, 2014, 2015; Karunanayaka et al., 2015). Experiences which are designed to create more intense and close interactions with the practitioners in context, while engaging them as co-participants in the process would be more desirable in promoting changes towards OEP.

Affordances of a DBR approach to support adoption of OEP
A DBR approach that involves researchers collaborating with practitioners in real-life settings to improve educational practices (Reeves, 2006), offers a feasible solution to ensure a sustainable and lasting impact on the adoption of OEP. DBR is a systematic but flexible methodology aimed at improving educational practices through iterative analysis, design, development and implementation in real-world settings (Wang and Hannafin, 2005). It comprises four phases: analysis of existing levels of expertise and practices by researchers and practitioners; designing, developing and implementing solutions as appropriate; testing and refinement of solutions in practice; and reflection by researchers and practitioners on authentic problems produce design principles and enhance solution implementation (Reeves, 2006). Being grounded in real-life context and conducted in collaboration with the practitioners, a DBR approach would have more potential in enacting desired changes of authentic educational practices.
DBR resembles action research (AR) in that it identifies real-world problems accompanied by subsequent actions to improve the status quo, and practitioners are highly involved in the research process. However, DBR is distinct from AR in two respects. Its major goal is to generate theory to solve authentic problems; and in it the roles of researchers and teachers as partners in the research process are central. In DBR, researchers take the initiative as both researchers and designers (Wang and Hannafin, 2005), whereas in AR, the practitioners initiate the research. Hence, when aiming at designing an approach to support changes in practices such as OEP among practitioners, DBR would be a more desirable approach.

“Design” is the key element of DBR. Yet, it extends beyond mere designing of interventions and testing them. Within a DBR framework, complex problems are addressed in real contexts in collaboration with practitioners; integrates known and hypothetical design principles with technological advances to render possible solutions to these problems; conducts rigorous and reflective enquiry to test and refine innovative learning environments as well as to define new design principles (Reeves, 2006) (see Figure 1).

The basic characteristics of DBR are defined as: pragmatic; grounded; interactive; iterative and flexible; integrative and contextual (Wang and Hannafin, 2005). With its ultimate goal of solving current real-world problems by designing and implementing interventions, while extending theories and refining design principles, rather than testing theory (The Design-Based Research (DBR) Collective, 2003), DBR emphasizes on an iterative cyclic process of designing-testing-refining solutions leading to a better understanding of the process. Being grounded in real-world context and conducted in collaboration with the practitioners, it has more potential in enacting desired changes of authentic educational practices. Further, given their joint role as designers and researchers, “[…] DBR researchers are not simply observing interactions, but are actually “causing” the very same interactions they are making the claims about […]” (Barab and Squire, 2004, p. 9). Such an iterative process can guide similar research and development endeavours (DBR Collective, 2003).

DBR is useful in designing solutions/strategies by researchers functioning as designers through an intervention, collaboratively with practitioners, to improve their educational practices. This provided the conceptual framework in guiding the design of an intervention in our study.

**Key research question**

The key research question of this study was as follows:

*RQ1.* How and to what extent a DBR approach impacted the adoption of OEP among teachers in terms of changes in their use of instructional resources, pedagogical thinking and pedagogical practices?

---

**Figure 1.**

Four phases of design-based research

Source: Adapted from Reeves (2006)
Methodology
The project implemented by OUSL involved the integration of OER and adoption of OEP by student teachers of the Faculty of Education, and ascertaining its impact on the use of instructional materials by teachers, and changes in pedagogical perspectives and practices of teachers. The focus of the current study was to explore how a DBR approach impacted on supporting OEP among the teachers.

Research design
The DBR (Reeves, 2006; Wang and Hannafin, 2005) approach adopted in this study involved researchers collaboratively working with practitioners in real-life settings to improve their practices along three aspects – instructional resource use; pedagogical perspectives; pedagogical practices.

The full programme of activities comprised the design and implementation of a professional development intervention programme for teachers on the integration of OER in teaching and learning, and ascertaining the impact of integrating OER and adoption of OEP in their teaching-learning process. This intervention programme was implemented in several stages, at nine OUSL centres representing nine provinces of the country.

Participants
Participants in the study comprised 230 student teachers during Stage 1, and 85 student teachers in 21 teams in Stage 2, representing the nine provinces of the country. A summary of the participant profile at Stage 1 is presented in Table I.

The participants constituted more females (66 per cent) than males (34 per cent), and all (100 per cent) were graduate teachers with more than half (57.8 per cent) being science graduates, and some (17.4 per cent) were with post graduate qualifications. A majority of the participants (75.7 per cent) were newcomers to the teaching profession with less than five years of teaching experience, and only very small percentage (1.7 per cent) with experience above 15 years. The distribution of participants at Stage 1 in the nine OUSL centres – Anuradhapura (A), Badulla (B), Batticaloa (Ba), Colombo (C), Jaffna (J), Kandy (Ka), Kurunegala (Ku), Matara (M) and Ratnapura (R) is presented in Figure 2.

The intervention framework
The intervention framework was designed as a contextualized, process oriented and a self-reflective enquiry following the four stages of the DBR approach – analysis, solution, testing and refinement and reflection (Reeves, 2006). This allowed examining the impact on the three aspects in focus – instructional resource use, pedagogical perspectives and pedagogical practices, through the provision of a carefully structured intervention process with specific strategies at each stage.

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Category</th>
<th>Number</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Female</td>
<td>152</td>
<td>66.1</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>78</td>
<td>33.9</td>
</tr>
<tr>
<td>Academic qualifications</td>
<td>BSc</td>
<td>133</td>
<td>57.8</td>
</tr>
<tr>
<td></td>
<td>BA/BCom</td>
<td>77</td>
<td>33.5</td>
</tr>
<tr>
<td></td>
<td>Others</td>
<td>20</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td>Post graduate</td>
<td>40</td>
<td>17.4</td>
</tr>
<tr>
<td>Professional experience</td>
<td>&lt; 5 years</td>
<td>174</td>
<td>75.7</td>
</tr>
<tr>
<td></td>
<td>6-15 years</td>
<td>52</td>
<td>22.6</td>
</tr>
<tr>
<td></td>
<td>&gt; 15 years</td>
<td>04</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Table I. Participant profile
A diagrammatic representation of the framework developed to design and implement the intervention using the DBR approach is presented in Figure 3.

The design of the intervention involved designing and developing a set of strategies based on existing theories and frameworks, tested design principles and prior research.

Strategies were developed to monitor and support teachers to gradually move from “no” or “low” usage of OER to “high” degrees of usage and creation of OER enhancing openness.

**Figure 2.**
Distribution of participants in the nine OUSL centres

**Figure 3.**
A DBR framework to provide experiences in OEP

<table>
<thead>
<tr>
<th>Aim – To support “Open Educational Practices” (OEP) among practitioners</th>
<th>DBR approach</th>
<th>Strategies/ experiences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>Instructional resource use (5R)</td>
<td>Pedagogical perspectives (5C)</td>
</tr>
<tr>
<td>Complex</td>
<td>Redistribute</td>
<td>Collaborative</td>
</tr>
<tr>
<td></td>
<td>Remix</td>
<td>Creative</td>
</tr>
<tr>
<td></td>
<td>Revise</td>
<td>Critical</td>
</tr>
<tr>
<td></td>
<td>Reuse</td>
<td>Challenging</td>
</tr>
<tr>
<td></td>
<td>Retain</td>
<td>Contextual</td>
</tr>
</tbody>
</table>

**Source:** Karunanayaka and Naidu (2016)
in their use of instructional resources, based on the 5R framework of OER – retain, reuse, revise, remix, and re-distribute (Wiley, 2014), and the OPAL framework to support OEP in two dimensions: openness in resource usage and creation vs openness in pedagogical models (Ehlers, 2011).

To enhance changes in teachers’ pedagogical perspectives and practices towards OEP, the experiences in the intervention were designed based on situated learning principles underlying a constructivist approach to learning (Brown et al., 1989; Duffy and Jonassen, 1991). These included specific strategies designed to support teachers moving from low towards high degrees of context centric, challenging, critical-thinking, creative, and collaborative thinking and practices (termed “5Cs”).

Specific strategies of the intervention in line with the four phases of DBR are presented in Table II.

As indicated in Table II, the four phases in the DBR approach were implemented in iterative cycles during the intervention, using different strategies. These strategies also served as a variety of data gathering methods, both qualitative and quantitative, systematically used at different stages of the intervention.

Initially, the current situation in relation to the three aspects – instructional resource use, pedagogical perspectives and pedagogical practices were analyzed, in collaboration with the practitioners. Next, as a solution, the intervention consisting of two key components – a series of interactive workshops and an enabling online environment in Moodle learning management system (LMS) was designed and implemented. Both these components included specific activities for capacity building, guiding, monitoring and supporting, as well as reviewing and evaluating, in relation to integration of OER by teachers. A variety of hands-on individual and group experiences were provided to engage teachers in the integration of OER in their teaching-learning process, while encouraging collaborative practices and promoting reflective practice.

Iterative cycles of testing and refinement of solutions in practice occurred during the series of pre, mid and post intervention workshops conducted at the nine OUSL centres, as well as constant interactions through the online environment in the LMS. Data gathering via multiple sources continued throughout, and these helped researchers to work together with practitioners in refining design strategies, based on contextual needs and complexities.

Throughout the intervention process, both teachers and researchers were encouraged to maintain reflective journals, writing their self-reflections on the experiences gained at different stages. Based on these reflections, teams of participant teachers and researchers compiled narratives in the form of “stories” at the end of the process. This helped in discussing and finding solutions to authentic problems, and to recommend effective design principles.

Methods of data collection and analysis

In keeping with the DBR process which involves integrative use of multiple data gathering methods over time, data were collected throughout the intervention process using questionnaire surveys, concept mapping, analyzing lesson plans, focus group interviews, records in the LMS, self-reflections and narratives (“stories”) of teachers. These strategies were employed at different stages – pre, mid and post intervention, and analyzed using qualitative and quantitative methods to capture changes in relation to adoption of OER and OEP in terms of changes in instructional resource usage, pedagogical practices and pedagogical perceptions of teachers.

Such methodological triangulation allowed providing a more comprehensive view of the whole, since each source of data gave a different view of the issues being studied (Morse, 1991). This process helped gaining a broader understanding of the effects of the intervention.

Interpretative phenomenological analysis (IPA) was the methodological construct used in the analysis of qualitative data that comprised content analysis and interpretation of
Table II. Strategies adopted during the intervention to promote adoption of OER and OEP

<table>
<thead>
<tr>
<th>Phase in DBR</th>
<th>Intervention stages</th>
<th>Specific strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis of practical problems by researchers and practitioners in collaboration</td>
<td>Pre-Intervention survey Orientation workshop</td>
<td>Reflecting on current thinking and practices in relation to use of instructional methods and materials by teachers Pre-Intervention questionnaire Individual concept mapping on “Openness in Education” Analyzing teachers’ lesson plans Focus group discussions Designed a sequence of experiences to enhance 5Rs in OER adoption, and 5Cs in teachers’ pedagogical thinking and practices Workshop activities LMS activities Monitoring activities</td>
</tr>
<tr>
<td>Development of solutions informed by existing design principles and technological innovations</td>
<td>Designing a series of interactive workshops Capacity building Monitoring/supporting Reviewing/evaluation Designing of a LMS to support OER adoption</td>
<td>Capacity building, monitoring and supporting teachers to adopt OER; Use of technology (LMS) to support OER adoption by teachers Designing a series of workshops to enhance 5Rs in OER adoption Workshop activities LMS activities Monitoring activities</td>
</tr>
<tr>
<td>Iterative cycles of testing and refinement of solutions in practice</td>
<td>Capacity building workshops LMS to support OER adoption Monitoring/supporting Workshops Reviewing/evaluation</td>
<td>Capacity building and guidance to adopt and integrate OER; lesson planning with OER integrations and upload in LMS; sharing OER in LMS; Monitoring, reviewing and supporting; enhance collaboration and cooperation; stimulate healthy competition; sharing of good practices; support extension activities; promote reflective practice; data collection Hands-on individual and group activities in identifying/searching/selecting OER and planning lessons with OER integrations; providing links to OER repositories via LMS; Encouraging teachers to use LMS to share OER found/reused/ revised/remixed/ created; providing constructive feedback; Constant communication via LMS; competition to find most active teacher and most active centre; encourage extension activities at school/centre/zonal levels; Concept mapping; reflective journal writing; questionnaire survey; focus group discussions</td>
</tr>
<tr>
<td>Reflection to produce design principles and enhance solution implementation</td>
<td>Teacher reflections researcher reflections</td>
<td>Use teachers and researcher’s reflections to find and implement solutions to authentic problems Compilation of “Stories” – by teachers and researchers, based on their reflections; Creation of a weblog to share the stories of their experiences</td>
</tr>
</tbody>
</table>
concept maps, open-ended questions in surveys, self-reflections, interview transcripts and narratives. The IPA strategy was used to explore in detail how individuals were perceiving the particular situations they faced and making sense of their personal and social world (Smith and Osborn, 2003), that helped discovering the meaning of experiences of participants through their and researchers’ interpretations by examining their “lived experiences” (Reid et al., 2005).

Further, such a “realist, process-oriented approach” that relied on understanding the processes by which a situation occurs via direct physical connection in the real world, offered a feasible approach for ascertaining causation, rather than a comparison of situations depending on the existence of the presumed cause (Maxwell, 2004; Mohr, 1999).

Results and discussion
Changes in instructional resources usage
Data obtained from questionnaire surveys revealed that at the pre-intervention stage, teachers’ use of OER was minimal, as indicated in Table III. Only 10 per cent have even heard the term “OER” and 3.9 per cent have used some OER materials, while use of online resources was only 20.4 per cent.

However, by mid-intervention this has significantly changed as shown by Figure 4, where a majority have claimed developing competencies in searching and identifying OER, identifying CC licence, adopting 5Rs, creating OER and integrating OER in their teaching practices.

Even though the number of participants decreased by the end of the intervention, many participants who remained were actively engaged not only in reusing OER, but also re-purposing them by translating into local languages, adapting to suit their contexts, and even creating OER on their own, as evident by the following excerpts:

When I use OER I modify it to local language. Some OER’s are advanced than I expect. Thus I edit it according to my lesson.

We were able to find interesting presentations on photosynthesis. We translated one presentation to Sinhala and used it to teach students. Sometimes we downloaded exercises and tests and made copies. Then we distributed among students.

The provisions of hands-on experience during the workshops and the Moodle LMS have vastly supported their use and adoption of OER:

Workshop activities helped us to identify relevant OER and identify the nature of their licenses […] it helped us to gain some knowledge and practice of the 4R concept through practical activities organized during the workshop.

We could also access the OER site created for us […] in the Moodle LMS and search for OER materials relevant to our subject areas. We could identify appropriate OER to integrate in lesson plans.

Hence, it was evident that starting from an initial state of “no or low-usage” of OER, the participants gradually moved towards adopting 5Rs and creating OER. This move from

<table>
<thead>
<tr>
<th>Aspect</th>
<th>Category</th>
<th>No.</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heard of OER before</td>
<td>Yes</td>
<td>23</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>207</td>
<td>90.0</td>
</tr>
<tr>
<td>Used OER before</td>
<td>Yes</td>
<td>9</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>221</td>
<td>96.1</td>
</tr>
<tr>
<td>Types of resources being used</td>
<td>Print</td>
<td>230</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>Audio/video/multimedia</td>
<td>162</td>
<td>70.5</td>
</tr>
<tr>
<td></td>
<td>Online resources</td>
<td>47</td>
<td>20.4</td>
</tr>
</tbody>
</table>

Table III. Teachers use of instructional materials at pre-intervention
“low” to “high” degrees of openness in two aspects – use, sharing and creation of OER, and innovative use of instructional resources, signifies enhanced diffusion of OEP (Ehlers, 2011). The specific strategies of the intervention – availability and accessibility of a variety of OER through the LMS that allowed 5Rs, and provision of hands-on activities to search, select and integrate OER in lesson plans have facilitated enhancing creativity and innovation in their use of resources.

### Changes in pedagogical perspectives

The structural and content analysis of the concept maps created by participant teachers at different stages of the intervention revealed the developments in teachers’ understandings and changing perspectives over time. The comparison of different versions of concept maps illustrated incremental changes in teachers’ perspectives, moving towards more “open” thinking, as depicted by Figures 5(a) and (b).

At the end of the process group concept maps were created by teams of participant teachers, elaborating their collective understandings and perceptions, as illustrated by Figure 6.

Figure 4.

<table>
<thead>
<tr>
<th>D 2.1 Searching for and identifying OER</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 5:</td>
</tr>
<tr>
<td>- 4:</td>
</tr>
<tr>
<td>- 3:</td>
</tr>
<tr>
<td>- 2:</td>
</tr>
<tr>
<td>- 1:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D 2.2 Identifying and understanding specific CC licenses of OER materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 5:</td>
</tr>
<tr>
<td>- 4:</td>
</tr>
<tr>
<td>- 3:</td>
</tr>
<tr>
<td>- 2:</td>
</tr>
<tr>
<td>- 1:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D 2.3 Reusing Revising Remixing Redistributing existing OER</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 5:</td>
</tr>
<tr>
<td>- 3:</td>
</tr>
<tr>
<td>- 2:</td>
</tr>
<tr>
<td>- 1:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D 2.4 Creating and uploading new OER materials into OER repositories</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 5:</td>
</tr>
<tr>
<td>- 4:</td>
</tr>
<tr>
<td>- 3:</td>
</tr>
<tr>
<td>- 2:</td>
</tr>
<tr>
<td>- 1:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D 2.5 Integrating OER into your teaching plans, lessons activities, assessments... etc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>- 5:</td>
</tr>
<tr>
<td>- 4:</td>
</tr>
<tr>
<td>- 3:</td>
</tr>
<tr>
<td>- 2:</td>
</tr>
<tr>
<td>- 1:</td>
</tr>
</tbody>
</table>
A designed based approach to OEP

Figure 5. Two versions of concept maps: pre intervention and mid intervention
Innovative, effective teaching and learning process
Enhanced student centered education
Lack of computers skills
Lack of language skills
Low internet facilities
Share OER among students
Mix two or more OER
Translate into local languages
Download and make copies
Integrate
Modify the content
No-cost access
User friendly
Easy to work
More information can be shared
Fewer restrictions
Anuradhapura group
13/06/2016
Reuse
Revise
Remix
Redistribute

Eg:

Attributes
Eg:

Benefits
Eg:

Teachers and students
Lack of computer skills
Lack of language skills

Simplified the content, reduced the content, expanded the content, changed content to match with the local context, translated into local languages, added content from other OER and media types, added newly created OER.

Facilitate new methods
Course design and development

Facilitate Open learning

Factors
Eg:

Materials
Teaching/Learning/research materials
Freely available
Open licensed

Eg:

OER repositories
Search engines

Subjects
Science
Mathematics
English
Biology
Chemistry
ICT
History
Physics

Presentation
Videos
Audios
Multimedia
Lesson plans
Class room activities
Animations
Graphics
Assessments
Software
Test
Text books
Modules
Full courses
Course materials

Eg:

Teaching/Learning/research materials
OER
Other OER
Creative works
Figure 6.
A group concept map – post-intervention
The final group concept maps constructed by teams of teachers demonstrated their enriched perspectives towards adoption of OER and OEP.

The concept mapping exercise has also supported teachers to reflect on changes in their own pedagogical thinking during the intervention process as claimed below:

The concept maps was a new experience […] which provided an easy way to summarize our findings throughout our activities and that new knowledge had made us special.

Even though teachers’ awareness on OER was minimal at the beginning, their perceptions on sharing resources was quite positive. However they were not very much concerned about copyrights when using resources. This view has changed during the process and teachers have started thinking more about using online resources legally and sharing resources, as indicated by the their reflections:

I have already used teaching materials which are available in the internet without checking the license agreement. But today I could get a real idea about the license agreement of the resources.

I can understand the meaning OER […]. Now I can use useful data and information legally. Before that I mostly use copyright data and information without permission […] now I use free usable data source with permission.

Further, it was evident that the participants’ pedagogical thinking has become more context centric, challenging, critical, and creative as well as collaborative, as following excerpts specify:

Through the integration of OER […] we have got the opportunity to “think out of the box” and release ourselves from the traditional teachers’ role. The creation of our own OER enhanced our thinking capabilities and contributed to the development of our personalities as well.

OERs helped us to plan and implement very attractive lessons. Here the teachers as well as students were encouraged to look for new knowledge. This created an opportunity for us to share subject related resources such as activities, assessments, video clips etc.

As evident by these quotes, engagement with OER has stimulated critical reflection among the teachers about their current practices and offered inspiration to attempt new practices with OER, and leading towards OEP (Beetham et al., 2012). The provision of specific “reflective” experiences in the intervention such as concept mapping, reflective journal writing and peer discussions, together with the individual and group activities on OER integration, and the enabling environment of the LMS to share OER, have significantly supported this. The compilation of “stories” based on their experiences with OER adoption process also promoted critical reflection and collaborative thinking leading to further enhancements in teachers’ pedagogical perspectives towards OEP.

Changes in pedagogical practices
Analysis of the lesson plans of teachers at the pre-intervention stage revealed that even though a majority (> 60-70 per cent) demonstrated use of a learner-centred pedagogic approach with learning activities in their overall approach to teaching, only a small percentage (> 10-20 per cent) demonstrated innovative learning designs, use of a variety of media types as learning resources, use of technology, providing opportunities for learner creativity, promoting self-regulated learning, linking with real-life situations and creating an enjoyable learning experience.

Yet, during the intervention, their pedagogical practices have notably changed, as revealed by the OER-integrated lesson plans, and also supported by their responses and reflections:

OER supported us to prepare a quality lesson plans for our lessons in all subjects. Then we were able to incorporate new techniques in our teaching-learning process.
Our lesson evaluators and colleagues highly appreciated OER incorporated lessons [...] It was very easy to get students’ attention and it supported to create a better learning environment.

Further, it was encouraging to observe many examples of sharing OER and disseminating the OER concept among others by the teachers via creation of OER and sharing them, conducting awareness raising workshops, publishing booklets etc:

I shared my OER knowledge with my school teachers by organizing a workshop. I created more than 30 OERs and uploaded to LMS and also searched more than 50 lessons to different subjects.

One member wrote a booklet in Sinhala about OER for teachers and any others who are interested in this concept.

We felt proud to publish a magazine on OER titled “Integrating OER in Learning Teaching Process.

Strategies such as motivating teachers to engage in lesson planning with OER integration and share them in the LMS, organizing a competition to identify the most active centre and the most active student teacher in each centre, have certainly impacted on promoting OER adoption among teachers. The design of such challenging activities related to their own contexts has stimulated teachers adopting critical, creative and collaborative approaches in their teaching practices. The shift from low to a high degree of OER use, and change of practices in the creative use of OER clearly suggests a move towards OEP. The diffusion of OEP was prominent along two dimensions – enhancements in the individual practices in innovative OER use as well as collaborative practices of sharing of resources, knowledge and practices (Ehlers, 2011), that has resulted due to the intervention.

Impact of the DBR approach in supporting OEP

The basic features of the DBR approach has been very supportive in enacting changes in educational practices of teachers towards OEP. In order to address a meaningful current problem faced by practitioners in relation to effective implementation of OEP, an educational intervention grounded in a robust theoretical framework was designed, developed and implemented in real-world settings, in collaboration with the practitioners (Reeves, 2006).

Situating the research work in naturalistic contexts enabled close interactions among researchers and practitioners, in their pursuit to find solutions to face the challenges in the adoption of OEP. The co-partnership link with the practitioners (teachers) was useful for researchers to test and refine design strategies in collaboration (Wang and Hannafin, 2005), according to the contextual and emerging needs during the process.

The iterative and flexible nature of the process allowed recognizing the complexities and dynamics of the real-world interactions and contextual limitations in the designs (Collins et al., 2004), and refining them as appropriate during the series of testing cycles, in the form of workshops and online activities via LMS.

The integrative use of multiple methods and a variety of data gathering strategies – both qualitative and quantitative, as required by the DBR approach enhanced the credibility of findings (Wang and Hannafin, 2005). It also generated a large amount of “thick” descriptive data sets (DBR Collective, 2003) which required systematic analysis and consensus building around interpretation of data to provide a comprehensive view of the process.

The armoury of tools and strategies created during the iterative process generates evidence-based claims about new design solutions, guidelines and frameworks to address the contemporary issues in integration of OER and OEP by practitioners, based on theoretical relationships. Thus, DBR approach, through its coherent methodology has enabled bridging theoretical research and educational practice (DBR Collective, 2003), and refining both theory and practice (Collins et al., 2004) through their synergy.
Conclusions and implications

It was evident that OEP among practitioners can be supported and “promoted” through a DBR approach using iterative analysis, design, development and implementation of carefully planned intervention strategies at different stages. Significant changes were observed in teachers’ use of instructional resources, their pedagogical thinking and pedagogical practices, due to this DBR intervention. The specific strategies designed and implemented in an intensive sequence of activities in the intervention during a series of workshops and provision of an enabling online environment supported both practitioners and researchers to gradually move through the four phases of DBR, analyzing issues, developing solutions, testing and refining solutions and reflecting on the experiences in their collaborative journey towards OEP.

The careful selection and design of activities to develop teachers’ competencies in identifying, using and creating OER has facilitated a shift in resource use from no or low usage of OER to reuse, revise and creation of OER. The opportunities provided to engage in sharing resources and challenging and innovative use of OER have enacted a change in the pedagogical perspectives and practices of teachers shifting from a content-centric and individualized pattern to more constructivist, context-centric and collaborative ways. The process of capturing teachers’ and researchers’ real-life experiences through a collection of “stories” around their experiences, enabled sharing as well as transfer of “good practices” in relation to adoption of OEP by practitioners.

The systematic and flexible methodology adopted via DBR by designing a framework aiming at improving educational practices was very useful to support changes in OEP among practitioners over time. This iterative process allowed the researchers to function as “designers”, while investigating real-life issues in collaboration with the practitioners through reflective enquiry to further refine innovative practices towards OEP. This provides valuable insights for improved design solutions for future interventions in similar contexts.

References


Creative Commons (n.d.), “About the licenses”, available at: http://creativecommons.org/licenses/ (accessed 2 January 2017).


Corresponding author
Shironica P. Karunanayaka can be contacted at: shironica.k@gmail.com
Learning analytics in higher education: an analysis of case studies

Billy Tak Ming Wong

University Research Centre, The Open University of Hong Kong, Hong Kong

Abstract

Purpose – The purpose of this paper is to present a systematic review of the mounting research work on learning analytics.

Design/methodology/approach – This study collects and summarizes information on the use of learning analytics. It identifies how learning analytics has been used in the higher education sector, and the expected benefits for higher education institutions. Empirical research and case studies on learning analytics were collected, and the details of the studies were categorized, including their objectives, approaches, and major outcomes.

Findings – The results show the benefits of learning analytics, which help institutions to utilize available data effectively in decision making. Learning analytics can facilitate evaluation of the effectiveness of pedagogies and instructional designs for improvement, and help to monitor closely students’ learning and persistence, predict students’ performance, detect undesirable learning behaviours and emotional states, and identify students at risk, for taking prompt follow-up action and providing proper assistance to students. It can also provide students with insightful data about their learning characteristics and patterns, which can make their learning experiences more personal and engaging, and promote their reflection and improvement.

Originality/value – Despite being increasingly adopted in higher education, the existing literature on learning analytics has focussed mainly on conventional face-to-face institutions, and has yet to adequately address the context of open and distance education. The findings of this study enable educational organizations and academics, especially those in open and distance institutions, to keep abreast of this emerging field and have a foundation for further exploration of this area.

Keywords Higher education, Learning analytics, ODL, Open and distance education

Introduction

Learning analytics (LA) refers to the process of collecting, evaluating, analysing, and reporting organizational data for decision making (Campbell and Oblinger, 2007). It involves the use of big data analysis for understanding and improving the performance of educational institutions in educational delivery. Open and distance learning (ODL) institutions present an ideal context for the use of LA as, with their large student numbers and the increasing use of the internet and mobile technologies, they already have a very substantial amount of data available for analysis with analytics.

Despite LA being increasingly applied in a wide range of educational organizations, the literature in this area has usually focussed on conventional face-to-face institutions. In the ODL setting, there is yet to be a systematic review summarizing existing work on the potential benefits of LA to open and distance institutions (Firat and Yuzer, 2016; Prinsloo and Slade, 2014), and relevant research findings potentially applicable to these institutions (Rienties et al., 2016).
This paper gives a systematic review of the mounting research work on LA that has been published in recent years to provide an overview of this emerging field and serves as a foundation for further exploration. It addresses the potential problems of ODL institutions that could be solved by using LA, and the benefits that could be obtained according to the existing case studies. It also presents a meta-analysis of relevant empirical studies which shows the effect of intervention for at-risk students based on the use of LA.

Related studies
LA involves the use of a broad range of data and techniques for analysis – covering, for example, statistical tests, explanatory and predictive models, and data visualization (Arroway et al., 2016). Various stakeholders, such as administrators, teaching staff, and students, can then act on the data-driven analysis. Without a standardized methodology, LA has been implemented using diverse approaches for various objectives. Gašević et al. (2016) summarized three major themes in LA implementation, namely, the development of predictors and indicators for various factors (e.g., academic performance, student engagement, and self-regulated learning skills); the use of visualizations to explore and interpret data and to prompt remedial actions; and the derivation of interventions to shape the learning environment. The diversity in LA implementation poses a challenge for education institutions which plan to be involved in it, leading to a commonly voiced question – “How do we start the process for the adoption of institutional learning analytics?” (Gašević et al., 2016, p. 4).

As an emerging field of study, an increasing number of case studies relevant to the implementation of LA in higher education have been published. However, only a small number of reviews summarize these individual case studies. Among them, Dyckhoff (2011) reviewed the research questions and methods of these studies. The findings showed that existing studies have focussed on six types of research questions: qualitative evaluation; quantitative measures of use and attendance; differentiation between groups of students; differentiation between learning offerings; data consolidation; and effectiveness. The research methods used include online surveys, log files, observations, group interviews, students’ class attendance, eye tracking, and the analysis of examination grades. Based on the results, suggestions were given on LA indicators for improving teaching.

Papamitsiou and Economides (2014) focussed on the impacts of LA and educational data mining on adaptive learning. They reviewed the experimental case studies between 2008 and 2013, and identified four distinct categories, namely, pedagogy-oriented issues, contextualization of learning, networked learning, and the handling of educational resources.

Also, Nunn et al. (2016) discussed LA’s methods, benefits, and challenges. It was found that the methods used included visual data analysis, social network analysis, semantic analysis, and educational data mining. The benefits of LA were seen to revolve around targeted course offerings; curriculum development; student learning outcomes; behaviours and processes; personalized learning; improvements in instructor performance; post-educational employment opportunities; and enhancement of educational research. The challenges included the tracking, collection, evaluation and analysis of data, as well as a lack of connection to learning science, the need for learning environment optimization, and issues concerning ethics and privacy.

Focussing on computer science courses, Ihantola et al. (2015) surveyed LA case studies in terms of their goals, approaches, contexts, subjects, tasks, data and collection, and methods of analysis. The goals were related to students, programming, and the learning environment. The approaches included case studies, constructive research, experimental studies, and survey research. They also found that most of the research work was undertaken in a course context, with the number of subjects ranging from 10 to 265,000, with 64 per cent of the studies having 500 or fewer subjects. In most of the studies, students
were required to complete multiple programming tasks. Over 60 per cent of the studies used automated data collection that logged students’ actions, and a variety of data analysis methods such as descriptive and inferential statistics.

The existing reviews of LA case studies provide a basic descriptive summary. However, as a new area in education, there remain many uncertainties for ODL institutions about involving themselves in it. To make an informed decision on whether or not to implement LA, a key question is: “What are the expected benefits for the institution?” This paper addresses this issue by surveying the outcomes of LA implementation for institutions.

Methodology
This study aims to investigate how LA has been used in higher education institutions and the outcomes obtained. Relevant case studies were collected from Scopus, using the key terms “academic analytics” and “learning analytics” for the period from 2007 to 2016. The studies were selected based on the following criteria:

(1) the study reported one or more empirical cases of the use of LA in a higher education institution;
(2) the institution in question was accredited by the government or government-related bodies;
(3) the institution had 1,000 or more students; and
(4) the source information contained the aims of using LA, a description of the analytics, its implementation and the outcomes.

An initial search returned 1,492 results. After screening, a total of 43 cases which fulfilled the criteria for inclusion were selected for further analysis. They were analysed in terms of their objectives, approaches, and major outcomes.

A meta-analysis was also conducted to synthesize the empirical findings reported in the case studies. Studies which included relevant quantitative data analysis were chosen, resulting in six studies on student support and analysis of learning behaviours, with the effect of LA intervention validated and reported.

Results
Benefits for institutions, staff, and students
A summary of the objectives and approaches of the use of LA in the institutions chosen is presented in Table AI. The benefits of LA for the institutions, staff and students revolve around the following aspects.

Improving student retention. Table I presents the use of LA which improved student retention. By closely monitoring students’ learning and persistence, undesirable learning behaviours and emotional states can be detected, and students who are at risk can be identified early. Factors leading to student dropout or retention can be identified and prediction models developed. Staff can take prompt follow-up action and provide proper assistance to students who need extra support, such as counselling, suggesting learning resources, and formulating individual learning plans. Students’ level of achievement, as well as their retention, can be enhanced.

Supporting informed decision making. Table II shows the use of LA which supported informed decision making. Institutions are provided with information and analyses generated from a massive amount of data for informed decision making. For example, planning can be carried out on course development and resources allocation on the basis of information about the popularity of courses, and types and frequency of materials reviewed by students.
Increasing cost-effectiveness. Table III presents cases of LA use which increased cost-effectiveness. LA can be integrated with other platforms such as the learning management system. Instructors can then access various kinds of information online for providing feedback and support to students. Analyses and feedback on students’ study progress can be delivered to staff, students, or parents in an automatic and cost-effective manner.

Understanding students’ learning behaviours. Table IV presents the use of LA for understanding students’ learning behaviours. By analysing diverse sources of data
(e.g. learning management systems and social networks), institutions and academic staff can understand the relationships among students’ utilization of resources, learning behaviours and characteristics, and learning outcomes, which helps them to evaluate the effectiveness of pedagogies and instructional designs for improvement. For instance, the use of LA helps to capture the students’ behaviours in watching course videos by highlighting the patterns of their preferences and behaviours as well as showing the parts of videos which were watched most and least frequently. Curriculum and learning materials can thus be better designed to address students’ preferences and needs.

Providing personalized assistance for students. Table V illustrates the use of LA for providing students with insightful data about their learning characteristics and patterns, which can make their learning experiences more personal and engaging, and facilitate their reflections and improvements while a course is still in progress. Early alerts can be automatically generated and sent to students if their academic performance is below a

<table>
<thead>
<tr>
<th>Institution</th>
<th>Major outcomes</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bridgewater College</td>
<td>Notifications were automatically generated and sent to students and their parents to recognize students’ good performance</td>
<td>Sclater et al. (2016)</td>
</tr>
<tr>
<td>Drexel University</td>
<td>Faculty, programme developers, and programme administrators were able to analyse the connections between a specific programme outcome and data related to that outcome</td>
<td>Harvey (2013)</td>
</tr>
<tr>
<td>Georgia Institute of Technology and Carnegie Mellon University</td>
<td>High reliability was achieved for analysing students' online discussion data</td>
<td>Wang et al. (2016)</td>
</tr>
<tr>
<td>Harvard University</td>
<td>A machine learning prediction model was shown to be effective for predicting students who would complete an online course</td>
<td>Robinson et al. (2016)</td>
</tr>
<tr>
<td>Lancaster University</td>
<td>Tutors could efficiently access various kinds of data for providing students with timely support</td>
<td>Sclater et al. (2016)</td>
</tr>
<tr>
<td>New York Institute of Technology</td>
<td>A dashboard simple and easy to use by staff was developed</td>
<td>Sclater et al. (2016)</td>
</tr>
<tr>
<td>Open University of Catalonia</td>
<td>Information could be updated and maintained automatically</td>
<td>Guitart et al. (2015)</td>
</tr>
<tr>
<td>Portland State University</td>
<td>Operation efficiency was increased, e.g. faster generation of reports</td>
<td>Blanton (2012)</td>
</tr>
<tr>
<td>Purdue University</td>
<td>Students who had engaged with the LA system sought more help and resources than other students</td>
<td>Arnold and Pistilli (2012)</td>
</tr>
<tr>
<td>Rio Salado College</td>
<td>The likelihood of successful course completion was accurately assessed</td>
<td>Smith et al. (2012)</td>
</tr>
<tr>
<td>The Hong Kong Institute of Education</td>
<td>There was greater interaction between teachers and students</td>
<td>Wong and Li (2016)</td>
</tr>
<tr>
<td>University of Adelaide</td>
<td>Lecturers were allowed to assess and monitor students’ collaboration in an online environment, without having to traverse a large discussion forum</td>
<td>Tarmazdi et al. (2015)</td>
</tr>
<tr>
<td>University of Michigan</td>
<td>The system demonstrated high scalability and extensibility</td>
<td>Mattingly et al. (2012)</td>
</tr>
<tr>
<td>University of Salamanca</td>
<td>The system allowed the provision of learning support to students in an automatic manner</td>
<td>Cruz-Benito et al. (2014)</td>
</tr>
<tr>
<td>University of the South Pacific</td>
<td>The utilization of open source resources could be modified and adapted by anyone to meet specific user needs</td>
<td>Prasad et al. (2016)</td>
</tr>
<tr>
<td>University of Sydney</td>
<td>LA features such as instant feedback and auto-grading are especially useful for instructors teaching subjects in computer science education</td>
<td>Gramoli et al. (2016)</td>
</tr>
</tbody>
</table>

Table III. Use of LA which increased cost-effectiveness
certain standard. Students can also be encouraged to engage more in the personalized learning activities which are conducive to success in their studies.

Timely feedback and intervention. Table VI presents the use of LA for timely feedback and intervention. Instructors can obtain up-to-date and holistic information about students’ study progress, so that timely feedback can be given and individualized interventions made. Students develop a sense of belonging to the learner community through personalized feedback given to them. For example, the use of social network analytics allows instructors to understand the development of the learner community and identify students who are

<table>
<thead>
<tr>
<th>Institution</th>
<th>Major outcomes</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ball State University</td>
<td>Data analyses showed the consistent predictive power of the LA system on students’ academic performance, persistence, retention and graduation</td>
<td>Jones and Woosley (2011)</td>
</tr>
<tr>
<td>Georgia Institute of Technology and Carnegie Mellon University</td>
<td>Students who displayed more higher-order thinking behaviours learnt more through deeper engagement with course materials displayed by their discussion behaviours These students in turn also learnt more than students who were constantly off topic in the forums Social-oriented topics triggered richer discussion compared with biopsychology oriented topics, and higher-order thinking behaviours tended to appear together within threads in the forums</td>
<td>Wang et al. (2016)</td>
</tr>
<tr>
<td>McGill University</td>
<td>It provides an unprecedented opportunity to use data from real learners in authentic learning situations to better understand learning processes The study demonstrated how to detect learner misconceptions Prediction precision and weighted relative accuracy were significantly increased</td>
<td>Poitras et al. (2016)</td>
</tr>
<tr>
<td>Oxford Brookes University</td>
<td>Problems were identified with ethnic minority students in particular courses</td>
<td>Sclater et al. (2016)</td>
</tr>
<tr>
<td>The Hong Kong Institute of Education</td>
<td>Potential indicators were found for predicting student performance, such as the contribution of in-depth contents in online discussion</td>
<td>Wong and Li (2016)</td>
</tr>
<tr>
<td>The Open University (UK)</td>
<td>Common pedagogical patterns were identified from learning designs, showing the relationship between learning activities and students’ learning outcomes</td>
<td>Toetenel and Rienties (2016)</td>
</tr>
<tr>
<td>The Technical University of Madrid</td>
<td>Relationship between student interaction and individual performance was identified</td>
<td>Fidalgo-Blanco et al. (2015)</td>
</tr>
<tr>
<td>The University of Melbourne</td>
<td>Relationships among students’ motivation, participation and performance in MOOCs were found</td>
<td>Barba et al. (2016)</td>
</tr>
<tr>
<td>The University of Melbourne</td>
<td>Learners’ learning progress could be visualized showing their development from novice to expert</td>
<td>Milligan (2015)</td>
</tr>
<tr>
<td>University of Adelaide</td>
<td>Lecturers could track the evolution of team roles across each study group and identify various sentiments within each group</td>
<td>Tarmazid et al. (2015)</td>
</tr>
<tr>
<td>University of Edinburgh</td>
<td>Patterns of students’ engagement in MOOC learning activities were found, showing differences in their learning behaviours between enrolments in the same courses</td>
<td>Kovanović et al. (2016)</td>
</tr>
<tr>
<td>University of North Bengal</td>
<td>Factors leading to students’ dropout were identified, such as pregnancy and the remoteness of residence locations</td>
<td>Yasmine (2013)</td>
</tr>
<tr>
<td>University of Rijeka</td>
<td>Student activities on the learning management system (e.g. assignment uploads and course views) were shown as predictors of academic success</td>
<td>Sisovic et al. (2015)</td>
</tr>
<tr>
<td>University of Santiago de Compostela</td>
<td>Teachers could understand more clearly how students behave during a course that facilitated the evaluation process</td>
<td>Gewerc et al. (2014)</td>
</tr>
</tbody>
</table>

Table IV. Use of LA which helped in understanding students’ learning behaviours
performing poorly or are isolated from the main discussion, and then provide intervention during discussion in real time. This is especially important for ODL institutions, where students may be using different study modes and social media is a major communication channel.

**Meta-analysis of the effect of interventions on student success**

An important function of LA is to predict at-risk students and deliver early alerts and interventions to them, in order to improve their academic attainment, and their retention and graduation rate. This section provides a meta-analysis of the various prediction models utilized in LA systems, and the effect of the intervention solutions on enhancing students’ success.

Among the case studies examined, only six which provided quantitative analysis results were selected and the results are synthesized in this section. The effect sizes for each analysis were calculated where the data required for the calculation were available, and a

<table>
<thead>
<tr>
<th>Institution</th>
<th>Major outcomes</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albany Technical College</td>
<td>Based on analysis of students’ study results, demographics and social data, at-risk students were identified for providing individual counselling</td>
<td>Karkhanis and Dumbre (2015)</td>
</tr>
<tr>
<td>Bridgewater College</td>
<td>Tutors were provided with detailed information to discuss with students on their progress against targets and suggested actions</td>
<td>Sclater et al. (2016)</td>
</tr>
<tr>
<td>Open Universities Australia The Technical University of Madrid</td>
<td>Students obtained from the system recommended content and activities and a personalized learning environment</td>
<td>Fidalgo-Blanco et al. (2015)</td>
</tr>
<tr>
<td>University of Michigan</td>
<td>Customized recommendations were provided, including suggestions on study habits, assignment practice, feedback on progress and encouragement</td>
<td>Mattingly et al. (2012)</td>
</tr>
<tr>
<td>Edith Cowan University</td>
<td>Students likely to need support were automatically identified and support staff could efficiently reach them for interventions</td>
<td>Sclater et al. (2016)</td>
</tr>
<tr>
<td>Marist College</td>
<td>Interventions resulted in a 6% improvement in final grades for the treatment group compared to the control group</td>
<td>Jayaprakash et al. (2014)</td>
</tr>
<tr>
<td>Northern Arizona University</td>
<td>Instructors’ feedback was available to individual students and to university personnel, facilitating a comprehensive support network for all students</td>
<td>Star and Collette (2010)</td>
</tr>
<tr>
<td>Purdue University</td>
<td>Interventions were provided to at-risk students, and a higher student retention rate was achieved</td>
<td>Arnold and Pistilli (2012)</td>
</tr>
<tr>
<td>San Diego State University</td>
<td>Interventions through e-mails were shown to be the best treatment within constraints, while having an impact on student achievement</td>
<td>Dodge et al. (2015)</td>
</tr>
<tr>
<td>University of Adelaide</td>
<td>The LA system allowed instructors to be aware when particular students are behaving differently from the others for making appropriate and timely interventions</td>
<td>Tarmazdi et al. (2015)</td>
</tr>
<tr>
<td>University of Edinburgh</td>
<td>Instant feedback was shown to be a useful LA feature for students in courses on computer programming</td>
<td>Kovavonci et al. (2016)</td>
</tr>
<tr>
<td>University of Michigan</td>
<td>Students were provided with feedback (e.g. grade prediction) for self-reflection</td>
<td>Mattingly et al. (2012)</td>
</tr>
<tr>
<td>University of Wollongong</td>
<td>Students who are isolated from the main discussion could be identified, and interventions could be provided during discussion in real time</td>
<td>Mat et al. (2013)</td>
</tr>
</tbody>
</table>
A descriptive comparison of the effect sizes across the studies was made. Table VII presents a summary of the predictive models and intervention solutions employed in the six case studies; and Table VIII summarizes the results of quantitative analyses for the intervention solutions and the effect sizes for each study.

To summarize, a common approach utilized in the cases of intervention for student success was to collect and analyse data from students’ learning activities and employ a specific computational model to predict and prioritize those students who were at-risk of dropping out or getting poor academic results. Based on the findings of the predictive modelling, subsequent measures can be taken for intervention. A common practice was to get academic staff to contact the at-risk students and provide personalized learning support to them. Such an approach to prediction and intervention was found to effectively enhance students’ success, as measured by various indicators such as GPA, study progress, the retention rate, and the graduation rate.

According to the meta-analysis of the quantitative results, all the institutions found improvement in the students’ success in the intervention group compared to the control group, although the effect size varied across different types of indicators for success and different institutions. For instance, the intervention groups in the case of Marist College showed a 6 per cent improvement in the students’ final grades compared to the non-intervention control groups (Sclater et al., 2016), while the effect size was in the range of small to medium based on Cohen’s (1988) convention. For the retention rate examined in Mattingly et al. (2012) for the Course Signal System of Purdue University, the intervention groups showed a nearly 50 per cent performance improvement compared to the control groups. In spite of the small sample size, the meta-analysis showed an encouraging result for the benefits of LA in aiding institutions to make effective informed decisions to improve students’ learning performance and success.

Discussion and conclusion

This study shows that positive outcomes have been widely reported in relevant case studies. The results suggest great potential for ODL institutions to utilize LA for analysing existing data, which is expected to benefit their operations in areas such as quality assurance and student support. This study also reviewed various predictive models for student success which were developed and validated to identify and prioritize students who may be in need of support. The quantitative analyses confirmed that the learning performance of these students improved after they had been approached for LA-based interventions. The findings of this study thus provide various stakeholders – institutions, staff, and students – with the benefits they may gain from LA.

In particular, the results related to student learning suggest that, to change students’ behaviours, it may suffice to simply make them aware of their learning engagement through LA tools in relation to other students or indicate that they are at risk (Jayaprakash et al., 2014; Sclater and Mullan, 2017). Complex data visualizations or dashboards may not be necessary. What is more important, as recommended in Gašević et al. (2016), is to help students to interpret correctly the information from visualizations or dashboards.

The meta-analysis revealed that only a few case studies related to LA implementation provided quantitative analyses data – a limitation which may be caused by the relatively new development of LA. Therefore, empirical investigations and validation of many new models and new theories in this area remain to be carried out. While an increase in the quantity of empirical and quantitative research can be expected in future, it is also important to develop and test innovative solutions supported by LA. Present LA-based interventions, as reviewed in this paper, were mostly based on the interaction and discussion between students and instructors. Although such interventions were shown to be effective in general, their effectiveness may vary among different groups of students in different contexts.
<table>
<thead>
<tr>
<th>Institution</th>
<th>Learning analytics system(s)</th>
<th>Predictive model</th>
<th>Intervention solution</th>
</tr>
</thead>
</table>
| Georgia Institute of Technology and Carnegie Mellon University (Wang et al., 2016) | Interactive-Constructive-Active-Passive (ICAP) framework | It was predicted that engaging in higher-order thinking behaviours results in better learning outcomes than paying general or focussed attention to course materials | Students’ online discussion behaviours were categorized into three types:  
Higher-order – the student has contributed at least one constructive or interactive post during a course  
Paying attention – the student has contributed at least one active post during the course but has not displayed any constructive or interactive posts  
No contribution to any on-topic discussion during the course  
Together with the students’ other persistent characteristics, treatment and control groups were formed to investigate differences in their learning outcomes  
Students’ posts in an online learning forum were extracted and analysed – how the students presented concepts, specifically whether they can make linkage among various concepts. Such a pattern was correlated with the grades they obtained. The findings can be used to guide interventions on students’ learning process, and inform ways to give feedback to improve teaching and learning |
<p>| Hong Kong Institute of Education (Wong and Li, 2016) | KeyGraph algorithm and Polaris (a software tool) | A test-mining analytical tool was used to predict students’ academic performance. The tool visualizes the hidden patterns and linkages among students’ learning activities. The findings of the study showed that this approach can provide insights into predicting students’ performance, and students with a higher grade tended to contribute more in-depth contents in an online learning environment | An online academic support environment was developed containing study skills materials and community support for specialists and student mentors. At-risk students identified by the predictive model were directed to the support environment |
| Marist College (Jayaprakash et al., 2014) | Open Academic Analytics Initiative | A machine learning algorithm and logistic regression were used to predict whether students are at risk based on their demographic details, aptitude data, and various aspects of their usage of the virtual learning environment obtained from the LA system | |</p>
<table>
<thead>
<tr>
<th>Institution</th>
<th>Learning analytics system ($)</th>
<th>Predictive model</th>
<th>Intervention solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nottingham Trent University</td>
<td>NTU Student Dashboard</td>
<td>Students’ engagement was assessed using indicators, such as door swipes into academic buildings, visits to the virtual learning environment, the submission of assignments, and the frequency of borrowing library resources. Each student received one of five engagement ratings: high, good, partial, low and not fully enrolled.</td>
<td>Tutors are prompted to contact students to give assistance when the students’ engagement drops off. Students can view their own engagement scores on the dashboard so that they will be self-motivated.</td>
</tr>
<tr>
<td>Paul Smith’s College</td>
<td>Rapid Insight’s Veera, Starfish EARLY ALERT, and CONNECT</td>
<td>Rapid Insight’s Veera combines different file types and uses automatic analyses and predictive modelling to identify at-risk students prior to their enrolment. Starfish EARLY ALERT automates data collection and uses analytics to increase the identification of at-risk students.</td>
<td>The Starfish EARLY ALERT and CONNECT automatically prioritize students who are identified as at-risk and facilitate intervention and outreach.</td>
</tr>
<tr>
<td>Purdue University</td>
<td>Course Signal System</td>
<td>The Course Signal System predicted students’ performance relying on a series of variables, including students’ demographic characteristics, academic performance, past academic history, and students’ efforts devoted to study.</td>
<td>Instructors provided real-time personalized feedback to each student based on the outcomes generated from LA, in which the student is informed about how he/she is doing.</td>
</tr>
<tr>
<td>Institution</td>
<td>Independent variable</td>
<td>Dependent variable</td>
<td>Statistical method</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>-----------------------------------------------------------</td>
<td>----------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Georgia Institute of Technology and Carnegie Mellon University</td>
<td>Higher-order thinking behaviours</td>
<td>Test score</td>
<td>Regression</td>
</tr>
<tr>
<td>Hong Kong Institute of Education</td>
<td>“Contribution” and “innovation” from students’ postings in discussion forum</td>
<td>Final grade</td>
<td>( \chi^2 ) test of independence</td>
</tr>
<tr>
<td>Marist College</td>
<td>Intervention</td>
<td>Final grade</td>
<td>One-way ANOVA</td>
</tr>
<tr>
<td>Nottingham Trent University</td>
<td>Level of engagement rating</td>
<td>Progression status</td>
<td>Descriptive categorical data analysis(^a)</td>
</tr>
<tr>
<td>Paul Smith’s College</td>
<td>Intervention</td>
<td>Grade, suspension or probation rate, graduation rate</td>
<td>Descriptive categorical data analysis(^a)</td>
</tr>
<tr>
<td>Purdue University</td>
<td>Intervention</td>
<td>Retention rate</td>
<td>( \chi^2 ) test of independence</td>
</tr>
</tbody>
</table>

Notes: \(^a\)The results presented in the case studies of these two institutions did not involve any statistical tests and complete information for the data — that is, sample size for each category was not provided. Therefore, no effect size could be calculated from the available data; \(^b\)the effect size was computed by combining the data for the second-year retention rate for three cohorts (2007, 2008, 2009) from the original tables in Mattingly et al. (2012).
A challenge in measuring the effectiveness of LA implementation lies in the difficulty of identifying the extent to which any change after the LA implementation is attributed to the LA itself. As discussed in Sclater and Mullan (2017), it may not be feasible to isolate the influence of LA when it is part of a wider initiative to develop data-informed approaches in an institution. The case studies published and reviewed in this paper would thus be biased to the institutions which only deployed LA without other measures in their data-informed approaches.

In the ODL context, work on LA remains at an initial stage. Features of ODL, such as open admission which allows a broad range of students to study the same course with very limited face-to-face interaction, are yet to be studied in relation to LA implementation. It is therefore suggested that future research can involve more fine-grained validation studies to identify the effect of the various factors involved the implementation of LA. In particular, investigation on those factors related to ODL institutions, staff and students, as well as the plausible constraints on their use of LA, would shed light on how they can benefit more from involvement in LA.

References


Further reading


Corresponding author

Billy Tak Ming Wong can be contacted at: tamiwong@ouhk.edu.hk
<table>
<thead>
<tr>
<th>Institution</th>
<th>Approaches</th>
<th>Objectives</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Albany Technical College</td>
<td>Monitoring, intervention</td>
<td>Identify at-risk students and provide them with counselling</td>
<td>Karkhanis and Dumbre (2015)</td>
</tr>
<tr>
<td>2. Ball State University</td>
<td>Monitoring, intervention</td>
<td>Identify at-risk students and provide them with counselling</td>
<td>Jones and Woosley (2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Increase effectiveness by reducing the time required to diagnose problems and targeting specific issues</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Help the institution to make informed decisions about student success programmes and retention services</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Allow students to become aware of the gaps between their behaviours and expected outcomes, to understand elements of their academic success, and to utilize on-campus resources to solve their problems</td>
<td></td>
</tr>
<tr>
<td>3. Bowie State University</td>
<td>Monitoring, intervention</td>
<td>Support student retention</td>
<td>Chacon et al. (2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Track students' progress towards graduation to facilitate decision making</td>
<td></td>
</tr>
<tr>
<td>4. Bridgewater College</td>
<td>Monitoring, intervention</td>
<td>Support students to do better than the national average</td>
<td>Sclater et al. (2016)</td>
</tr>
<tr>
<td>5. California State University</td>
<td>Monitoring</td>
<td>Analyse how students use the learning management system</td>
<td>Allen et al. (2012)</td>
</tr>
<tr>
<td>6. Drexel University</td>
<td>Updating data and curriculum</td>
<td>Measure the effectiveness of specific course components through maintaining data records aligned with the curriculum, courses and syllabi, course learning objectives and assessment strategies</td>
<td>Harvey (2013)</td>
</tr>
<tr>
<td>7. Edith Cowan University</td>
<td>Monitoring, intervention</td>
<td>Manage student learning outcomes and performance criteria</td>
<td>Sclater et al. (2016)</td>
</tr>
<tr>
<td>8. Georgia Institute of Technology and Carnegie Mellon University</td>
<td>Monitoring, analysis</td>
<td>Better scaffolded online discussion to improve learning in a MOOC context</td>
<td>Wang et al. (2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Explore effects of higher-order thinking behaviours in learning</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identify kinds of discussion behaviours associated with learning</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigate types of learning materials which trigger richer discussion</td>
<td></td>
</tr>
<tr>
<td>9. Harvard University</td>
<td>Monitoring, prediction</td>
<td>Analyse the extent to which students' responses about motivation and utility value can predict persistence and completion of study</td>
<td>Robinson et al. (2016)</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Institution</th>
<th>Approaches</th>
<th>Objectives</th>
<th>Source</th>
</tr>
</thead>
</table>
| 10. Lancaster University          | Monitoring, intervention, feedback | Allow tutors to access the transcripts of their students  
Ensure student work is graded and feedback given to students in a timely manner | Sclater et al. (2016)      |
| 11. Loughborough University       | Feedback                    | Provide academics with a better and more holistic picture of student engagement  
Provide staff with actionable insights into student learning experience  
Provide students with their own educational data in a meaningful way | Sclater et al. (2016)      |
| 12. Manchester Metropolitan University | Monitoring, curriculum design | Improve student experience as reflected in the National Student Survey  
Provide data for improving the undergraduate curriculum | Sclater et al. (2016)      |
| 13. Marist College                | Prediction, intervention    | Predict academic success  
Provide interventions  
Identify misconceptions of medical students as reflected in their interactions in the online learning environment | Jayaprakash et al. (2014)  |
| 14. McGill University             | Monitoring, analysis        | Create an at-risk model to identify students in need of support  
Improve student retention in their first year of study  
Provide information that could support counsellor in their work | Poitras et al. (2016)      |
| 15. New York Institute of Technology | Prediction, intervention    | Facilitate online interaction between students and instructors  
Allow students to receive direct feedback on issues such as academic concerns and grades | Sclater et al. (2016)      |
| 16. Northern Arizona University   | Feedback                    | Enhance retention and improve attainment  
Increase students’ sense of belonging within the course community, particularly with tutors  
Identify at-risk students  
Suggest alternative modules to students which are more appropriate for their needs  
Identify automatically pieces of knowledge taught in each subject  
Gather students’ information  
Keep information updated  
Improve student experience  
Support progress evaluation of modules and programmes, and the identification of priorities at an institutional level | Star and Collette (2010)  
Atif et al. (2013)  
Guitart et al. (2015)  
Sclater et al. (2016)  |

Table AI. Learning analytics in higher education (continued)
<table>
<thead>
<tr>
<th>Institution</th>
<th>Approaches</th>
<th>Objectives</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>21. Paul Smith’s College</td>
<td>Monitoring, intervention</td>
<td>Identify at-risk students and prioritize outreach for them</td>
<td>McAleese and Taylor (2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Provide more efficient and effective interventions for student success</td>
<td></td>
</tr>
<tr>
<td>22. Portland State University</td>
<td>Information management</td>
<td>Make information more accessible and easier to use</td>
<td>Blanton (2012)</td>
</tr>
<tr>
<td>23. Purdue University</td>
<td>Monitoring, intervention</td>
<td>Give students early and frequent performance notifications</td>
<td>Arnold and Pistilli (2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Help faculty members to steer students towards additional campus resources as needed</td>
<td></td>
</tr>
<tr>
<td>24. Rio Salado College</td>
<td>Prediction</td>
<td>Identify factors having a significant statistical correlations with final course outcomes</td>
<td>Grush (2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Discover approaches that could be applied with minimal support and are scalable to a large number of courses</td>
<td></td>
</tr>
<tr>
<td>25. San Diego State University</td>
<td>Intervention</td>
<td>Identify methods and interventions that would alleviate students' failure</td>
<td>Dodge et al. (2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Help students towards additional campus resources as needed</td>
<td></td>
</tr>
<tr>
<td>26. The Hong Kong Institute of Education</td>
<td>Monitoring, feedback</td>
<td>Provide insights into predicting students’ performance</td>
<td>Wong and Li (2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Develop measures to assess students’ online learning</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Boost teachers’ and students’ interaction</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Allow students to realize their knowledge discovery</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Facilitate teachers to assess students’ performance</td>
<td></td>
</tr>
<tr>
<td>27. The Open University (UK)</td>
<td>Monitoring, intervention, personalization</td>
<td>Identify learners at risk and needing support</td>
<td>Rienties et al. (2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve learning design</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deliver personalized intervention for students</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Achieve cost-effectiveness</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Identifying patterns</td>
<td>Identify common patterns in course design</td>
<td>Toetenel and Rienties (2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Find out pedagogical implications for various patterns and learning designs</td>
<td></td>
</tr>
<tr>
<td>29. The University of Adelaide</td>
<td>Monitoring, feedback</td>
<td>Analyse students’ online discussion data, such as team mood, role distribution and emotional climate</td>
<td>Tarmazdi et al. (2015)</td>
</tr>
<tr>
<td>30. The University of East London</td>
<td>Monitoring, feedback</td>
<td>Develop students’ soft skills necessary for collaborative work</td>
<td>Schleyer et al. (2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Monitor student attendance and learning activities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Collect student data, such as demographic information, library activities, coursework, and download of free books</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Send automated e-mails to students showing their attendance, and warnings to students without satisfactory attendance</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Institution</th>
<th>Approaches</th>
<th>Objectives</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>31. The University of Melbourne</td>
<td>Monitoring, analysis</td>
<td>Investigate how motivation and participation influence students’ performance in a MOOC</td>
<td>Barba et al. (2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Analyse how MOOC participants use online forums to support learning</td>
<td>Milligan (2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigate how students interpret feedback delivered via learning analytics dashboard and the relevant influence on their learning strategies and motivation</td>
<td>Corrin and Barba (2015)</td>
</tr>
<tr>
<td>32. Universidad a Distancia de Madrid</td>
<td>Monitoring, analysis</td>
<td>Find predictors of teamwork and commitment as cross-curricular competences</td>
<td>Iglesias-Pradas et al. (2015)</td>
</tr>
<tr>
<td>33. University of Edinburgh</td>
<td>Analysis, prediction</td>
<td>Examine MOOC data about students who enrolled in the same course at least twice</td>
<td>Kovanović et al. (2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identify changes in their behaviours between the two enrolments to the same course</td>
<td></td>
</tr>
<tr>
<td>34. University of Maryland, Baltimore County</td>
<td>Monitoring, feedback, reflection</td>
<td>Reduce student barriers</td>
<td>Mattingly et al. (2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Create a community of learners</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve students’ self-awareness by providing feedback</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Provide early alerts to students if their GPA falls below a level</td>
<td></td>
</tr>
<tr>
<td>35. University of Michigan</td>
<td>Monitoring, personalization, reflection</td>
<td>Identify at-risk students</td>
<td>Mattingly et al. (2012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Provide personalized feedback to students</td>
<td></td>
</tr>
<tr>
<td>36. University of New England</td>
<td>Monitoring, intervention</td>
<td>Foster a sense of community among students studying part-time, at a distance as well as on-campus</td>
<td>Sclater et al. (2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identify students who are struggling in order to provide timely support</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Develop a dynamic, systematic and automated process to capture the learning well-being status of students</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Encourage peer-to-peer student networking</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Disseminate information and connect support staff with the students</td>
<td></td>
</tr>
<tr>
<td>37. University of North Bengal</td>
<td>Prediction</td>
<td>Examine the predictive relationship between learners’ pre-entry demographic information and their dropout behaviours</td>
<td>Yasmine (2013)</td>
</tr>
<tr>
<td>38. University of Rijeka</td>
<td>Data mining, analysis</td>
<td>Find out factors leading to student success in study</td>
<td>Sisovic et al. (2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identify problems timely and increase the course pass rate</td>
<td></td>
</tr>
</tbody>
</table>

Table A1.

(continued)
<table>
<thead>
<tr>
<th>Institution</th>
<th>Approaches</th>
<th>Objectives</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>39. University of Salamanca</td>
<td>Information extraction, analysis</td>
<td>Extract information useful for teaching/administrative staff, such as interaction of students with peers, teachers, the system, and course contents  Provide teachers with tools to facilitate managerial tasks</td>
<td>Conde et al. (2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Support practical learning in a 3D virtual environment, analyse the problems that arisen, and report relevant data to students and teachers</td>
<td></td>
</tr>
<tr>
<td>40. University of Santiago de</td>
<td>Analysis, evaluation</td>
<td>Generate automatically reports of learners’ activities that take place in a virtual learning environment  Improve the efficiency of the evaluation process</td>
<td>Gewerc et al. (2014)</td>
</tr>
<tr>
<td>Compostela</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41. University of Sydney</td>
<td>Analysis, observation</td>
<td>Identify the relationship among student performance, choices of programming languages for study, and times at which a student starts and stops working on an assignment</td>
<td>Gramoli et al. (2016)</td>
</tr>
<tr>
<td>42. University of the South Pacific</td>
<td>Monitoring</td>
<td>Track individual learners’ online and offline interactions with open learning resources</td>
<td>Prasad et al. (2016)</td>
</tr>
<tr>
<td>43. University of Wollongong</td>
<td>Analysis, intervention, reflection</td>
<td>Visualize patterns of student interactions on discussion forums  Allow instructors to identify at-risk students and potentially high and low performing students for planning interventions, and the extent to which a learner community is developing in a class</td>
<td>Mat et al. (2013)</td>
</tr>
</tbody>
</table>
Student characteristics and knowledge on ODL concepts at first registration
A case study from OUSL

K.A.J.M. Kuruppuarachchi and K.O.L.C. Karunanayake
Department of Botany, The Open University of Sri Lanka, Nawala, Sri Lanka

Abstract
Purpose – The purpose of this paper is to identify socio-economic/demographic characteristics and to evaluate the knowledge on different open distance learning (ODL) concepts of BSc undergraduates of The Open University of Sri Lanka (OUSL) at first registration.
Design/methodology/approach – The surveying technique was adapted with a piloted structured questionnaire consisting of two components. The structured component was used to evaluate personal, socio-economic and demographic data. The open ended component evaluated the student’s perception on ODL concepts. The questionnaire was randomly adapted to 456 (35 percent Colombo Regional Centre (CRC) registrants) prospective BSc undergraduates at first registration time at the CRC in 2014. Data collected from the structured component were frequency tabulated and cross-tabulated with the SPSS computer software. Responses of the open ended part were examined, categorized and the frequency percentages of each response category were calculated.
Findings – The structured component recognized that the majority of BSc undergraduates of the OUSL represent employed (53 percent), late adolescents (92 percent below age 27) who reside in rural or semi-urban areas (75 percent). They belong mostly to the lower middle class and 69 percent are from families which have a monthly family income below SLR30,000/(USD208). Answers of the open ended component on ODL concepts recognized that, prior knowledge on ODL concepts were developed by most BSc undergraduates. Approximately 50 percent of respondents perceived OUSL as an institute which facilitates working people by conducting part time-based or distance mode education with self-learning features. In total, 56.9 percent students perceived the role of an ODL teacher correctly as a facilitator or a guide. The educational process was perceived correctly as an ODL system by 52 percent, while the remainder also identified the system to be a more self-study and student centered flexible learning system. However, the role of a BSc student at OUSL was recognized as self-independent learners by only 36.7 percent and the majority had no clear perception of the role they have to play as an ODL student. Hence, more attention should be paid to make students recognize the role they have to play in an ODL system in order to succeed at OUSL.
Originality/value – Although research has been carried out periodically on the process of ODL education system at OUSL, on the graduate (output) and dropouts, etc., not many have focused on the nature of input such as characteristic features of first registrant and their prior knowledge on ODL. As the output invariably depends on the input and the process, this type of survey is timely and novel.
Keywords Open and distance education (ODE), BSc undergraduate, First registrant, ODL concepts, OUSL, Teaching methodology
Paper type Case study
Introduction
Teaching at a distance is characterized by separation of teacher and learner and learner from the learning group, by replacing interpersonal face-to-face communication of conventional education by an impersonal mode of communication mediated by technology (Keegan, 1996). At present, distance education systems and open education have become a feasible and viable alternative to those who were denied education due to one reason or the other (Nigam and Joshi, 2007). Although subjects dealing with humanities and social sciences have been offered through open distance learning (ODL) in the absence of major constraints, teaching science and technology courses through ODL is more challenging, mainly due to intensive practical work (Nigam and Joshi, 2007). Student dropout rates in ODL higher education systems are typically higher than in face-to-face conventional systems of higher education (Barefoot, 2004), and a study focusing dropouts of the BSc degree program at the Indira Gandhi National Open University, India (Fozdar et al., 2006) has come out with several significant issues faced by students. The lack of time (due to employment, family commitments), distance to study centers, absence of interaction with students, insufficient academic support, expectations of the program not being met and financial constraints were some of the key issues identified in this study. Investigations by Aluwihare and De Silva (2016) demonstrate that similar reasons contribute toward the prolonged time duration taken toward completing the Engineering degrees at the Open University of Sri Lanka (OUSL).

In the Sri Lankan context, OUSL is the premier and only state ODL institution where students can pursue their studies through ODL methodology. Established in 1980 following the Open University UK concept, the institute has the same legal and academic status as any other national university in Sri Lanka. Due to the nature of its teaching methodology and infrastructure, the OUSL is able to serve a large student population spread throughout the country. Data for 2015 indicate that it caters to 42,911 students, in either the Central campus in Nawala (Colombo Regional center) or in any one of the other seven regional centers in Kandy, Matara, Jaffna, Anuradhapura, Badulla, Kurunegala or Batticaloa. The University has five faculties: Natural Sciences, Health Sciences, Engineering Technology, Humanities and Social Sciences and Education.

The Faculty of Natural Sciences conducts several study programs such as short courses, diploma programs, certificate programs, etc. However, the Faculty’s primary concern is the degree program leading to the BSc degree. As at 2015, 4,261 students were registered for and following the BSc degree program at OUSL and approximately 1,300 students enroll annually.

Unlike in the other state universities, the students enrolling to follow degree programs at OUSL are comparatively diverse (Jayathilake et al., 1997). Tantrigoda et al. (2014) state that as the majority of students enrolling for the BSc degree are adolescents who have undergone formal education, shifting to distance education with self-learning maybe a totally new experience for them. Students sometimes opt for ODL expecting it to be easy (Carnevale, 2000) and often face huge difficulties when they realize ODL requires equal if not more effort than conventional education (Fozdar and Kumar, 2006). Further, Kuruppuarachchi et al. (2016) ascertain that, the components of the education processing system (PS) of the OUSL play a comparatively different role to that of a conventional university and that the input (students enrolling) and process are comparatively different in ODL, but the quality of output (the graduate) should be equal or higher than that of a conventional setup. Evaluation of the nature of input of PS is very important for future educational strategic changes and for efficient quality products Kuruppuarachchi et al. (2016). Due to these reasons the present study was carried out with 456 (35 percent) prospective BSc undergraduates of OUSL at first registration at the Colombo Regional Centre (CRC) in the 2013/2014 academic year. The total number of new registrants for the BSc degree program at OUSL was 2,373 in the 2013/2014 academic year. Of this 55 percent
registered at CRC. The survey explores the existing knowledge on ODL concepts of these prospective students while also attempting to get an insight into the socio-economic background characteristics of these students, in the view that the findings of the study will help ODL managers/administrators to make future direction of the BSc degree program.

Methodology
The surveying technique was adapted with a piloted structured questionnaire. The questionnaire consisted of two components. The first (structured) component: with 14 sections was used to evaluate personal, socio-economic and demographic data (Kuruppuarachchi and Gunerathne, 2014), while the open ended component facilitated the recognition of student’s existing perceptions on varying ODL concepts. The questionnaire was randomly distributed at the CRC (Central campus) and collected from 456 (35 percent of students registered at CRC) prospective BSc undergraduates of OUSL at first registration for the academic year 2013/2014. The statistics for BSc student enrollment clearly shows that all other six regional centers and 18 study centers belonging to OUSL collectively represent only 45 percent of the new registrant population. As 55 percent of new registrants are at CRC, it seems justifiable to collect data from CRC and generalize results of the study for BSc undergraduates of OUSL.

Statistical analyses
Quantification of responses of the structured component. Data of responses of the first component (structured component) of the questionnaire were used to make a frequency tabulation test with the SPSS computer package.

Quantification of responses of the open ended component. The open ended questions received diverse as well as similar responses. Some seemingly different responses tended to have similar meanings. All responses were carefully examined and then the main categories were chosen. Each individual response was analyzed and placed in the relevant category. Quantification of each similar category of responses of the open ended part was carried out by calculating frequency percentages of each response category as follows (Kuruppuarachchi and Gunerathne, 2014; American Association for Public Opinion Research, 2000):

\[
\text{Frequency percentage of "X" - category of responses} = \frac{\text{Number of responses belong to "X" category}}{\text{Total number of responses for the open ended part}} \times 100
\]

The open ended part of the questionnaire was adapted to explore the prior knowledge of BSc undergraduate of OUSL on “different ODL concepts.” This consisted of four parts. These are: teaching-learning process of OUSL; main instructional media used in OUSL; the role of the teacher of OUSL; the future role as an OUSL student. The responses provided were scrutinized with an open mind, without any restrictions, discrimination or preformed perceptions. The responses of students for each part were categorized, then quantified and are presented as described under the result and discussion section.

Results and discussion
Structured component
The results of the structured component of the questionnaire describe the socio-economic characteristics of BSc first registrants of OUSL.
The highest educational qualification of parents or any other family member of most registrants was found to be (ca. 84 percent) below G.C.E. (A/L). This implies that the majority of parents are middle class employees. Further, the permanent residences of the new registrant to the BSc degree are mainly rural or semi-urban (74.4 percent). The monthly mean household income in year 2012 of Sri Lankan people was recognized as Urban (USD471), rural (USD291), estate sector (USD220), respectively (Department of Census and Statistics, 2013). Result of the study show that the majority of students registered for the BSc degree program in 2013/2014 academic year belong to the low income category. That is, the total monthly family income is less than SLR30,000/-(USD208) for 68.8 percent (Figure 1). The government policy of Sri Lanka after independence in 1948 was to provide free school and university education. However, after 2010 this policy was changed somewhat to allow recognized private sector and international universities to established and award degrees at a fee for those who wished to obtain them. Charges of such institutes are comparatively high. Thus, attraction of lower income categories to OUSL would be due to the affordable and easy payment scheme provided by OUSL (total Ca. SLR150,000.00 for 108 credits for the BSc degree). Similar results were seen in the Bangladesh Open University, where 58.3 percent of the learners belonged to the lower-middle-income group population (Anwarul et al., 2010).

Most of the students that registered for the BSc degree are unmarried (92 percent), employed (53 percent) (Figure 2) and belong to the late adolescent (48 percent age 18-22, 92 percent below 27 years) age category. The age at graduation of students of OUSL in 1996 showed that the majority of BSc graduates were less than 30 years (76 percent), Bachelor of Engineering Technology (BTech) graduates 31-41 years and that 58 percent of LLB (law) graduates ranged from 31-71 (Senarathne et al., 2001). As such it is clear that BSc graduates have been comparatively young even in the past. However, in the
Open University of UK, the majority of undergraduates were aged between 31 and 60 and the average age was 43 years (Feinstein et al., 2007). However, a study carried out in 1997 in Sri Lanka (Jayathilake et al., 1997), to recognize characteristic features of OUSL students in comparison with conventional university students proved different. Our study in 2013 results in similar findings for students’ characteristics but attraction of employed students to OUSL has increased compared to 1997. This is a good indication from the ODL concept point of view. However, one has to bear in mind that now there are more private universities. Some of these provide part time-based education; hence the competition faced by OUSL is higher. If the BSc program of OUSL continues to attract a majority of lower middle class or low income groups into the system, it will affect the students’ intra-social interactions as well as the sustainability of the institution in the long run. Thus, the administration of the university should attempt to attract clients of other social strata as well. Most employment opportunities in Sri Lanka belong to the private sector and usually they consider the social strata, school, etc. of the candidate. If the OUSL is recognized as a university which caters mostly to low income social strata, it will affect the employability of OUSL, BSc graduates in the long run. In addition, retaining learners has always been an ongoing challenge for ODL institutions worldwide (Fozdar et al., 2006; Bandarage, 2016). According to OUSL statistics, 15 percent ±7 of new registrants of BSc program dropout without starting the academic activities as “Non-starters” (Bandarage, 2016). The student dropout rate has been high (ca.50 percent) in first and second years of BTech. study programs at OUSL (Dadigamuwa and Senanayake, 2012). It appears that a considerable percentage of undergraduates of OUSL give up their degree program without completion. One reason for this could be the “poor socio-economical background” of BSc undergraduates.

Open ended component
The primary purpose of this component of the questionnaire was to evaluate the existing knowledge on “ODL concepts” of BSc undergraduates of OUSL at initial registration. The open ended part consisted of four questions as: what is the teaching-learning process of OUSL; the main instructional media adapted by OUSL; the role of a teacher of OUSL; the role of a student of OUSL. The responses provided by students for each of these questions are tabulated from Tables I-IV, respectively.

<table>
<thead>
<tr>
<th>Response number</th>
<th>Response category (example of answers given)</th>
<th>Response frequency %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Open distance system (does not strictly follow the basic qualification for entry as a state university, any person can enter, not a full time program, most activities are conducted during weekends); does not have a fixed age limit; can follow the program while working</td>
<td>51.97% (237)</td>
</tr>
<tr>
<td>2.</td>
<td>Not answered for the question</td>
<td>17.76% (81)</td>
</tr>
<tr>
<td>3.</td>
<td>Self-study system (students have to study and learn concepts themselves, students are self-motivated, academic achievements depend on self-ability and efficiency)</td>
<td>12.06% (55)</td>
</tr>
<tr>
<td>4.</td>
<td>Flexible learning system (a student can select the amount of workload favorable to him)</td>
<td>5.04% (23)</td>
</tr>
<tr>
<td>5.</td>
<td>Conducting ODL system of OUSL is more student centered. Different media used to facilitate students. Selection of subject units and amount is a student’s own choice</td>
<td>4.16% (19)</td>
</tr>
<tr>
<td>6.</td>
<td>No idea about the teaching-learning process at OUSL</td>
<td>5.7% (26)</td>
</tr>
<tr>
<td>7.</td>
<td>OUSL is conducting formal lectures, teaching system similar to other conventional universities</td>
<td>1.53% (67)</td>
</tr>
<tr>
<td>8.</td>
<td>Teaching-learning process is strictly part time</td>
<td>1.75% (68)</td>
</tr>
</tbody>
</table>

Table I. BSc new registrants’ perception of the teaching-learning process adapted by OUSL
The responses for the “nature of teaching-learning process of OUSL” clearly show that students have sufficient knowledge on this concept. For example, they have recognized that the teaching-learning process as an open distance system (51.97 percent), self-study system (12.06 percent), flexible learning (5.04 percent), and students centered (4.16 percent), respectively (Table I). Only, a very low amount of response frequencies (1.53 percent) were allocated as “OUSL is conducting formal lectures” similar to a conventional setup.

Aluwihare and De Silva (2016) state that about 86 percent BTech undergraduates of OUSL did not show sufficient knowledge on the open and distant teaching-learning process and concepts. Further, this was found to be a contributory factor in prolonging the time

<table>
<thead>
<tr>
<th>Response number</th>
<th>Response category</th>
<th>Response frequency %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mainly printed materials (developed by OUSL, and supplementary material)</td>
<td>45.78% (174)</td>
</tr>
<tr>
<td>2</td>
<td>Combination of different instruction media (printed material + educational video + internet, etc.)</td>
<td>35.26% (134)</td>
</tr>
<tr>
<td>3</td>
<td>Mainly internet-based education</td>
<td>9.73% (37)</td>
</tr>
<tr>
<td>4</td>
<td>Not answered the question</td>
<td>8.68% (33)</td>
</tr>
<tr>
<td>5</td>
<td>Mainly educational video and electronic media</td>
<td>0.52% (2)</td>
</tr>
</tbody>
</table>

Table II. BSc new registrants’ perception of the main instructional media used in OUSL

<table>
<thead>
<tr>
<th>Response number</th>
<th>Response category</th>
<th>Response frequency %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>OUSL teacher recognized as “a helper for studies,” “a good counselor,” “a facilitator” “or a good advisor”</td>
<td>56.9% (185)</td>
</tr>
<tr>
<td>2</td>
<td>Formal lecturer role: teaching subject matter like in our school classroom, formal teacher role, should be a lecturer like in other government universities</td>
<td>20.3% (66)</td>
</tr>
<tr>
<td>3</td>
<td>Not answered the question</td>
<td>17.23% (56)</td>
</tr>
<tr>
<td>4</td>
<td>No Idea on the role of the OUSL teacher</td>
<td>2.76% (9)</td>
</tr>
<tr>
<td>5</td>
<td>Teacher should be a very friendly, kind hearted person, should able to make interactions as well as emotional attachments with students</td>
<td>2.46% (8)</td>
</tr>
<tr>
<td>6</td>
<td>OUSL teacher should play multiple roles as a good teacher, guide, helper, etc.</td>
<td>0.3% (1)</td>
</tr>
</tbody>
</table>

Table III. The role of the teacher at OUSL: as perceived by BSc new registrants

<table>
<thead>
<tr>
<th>Response number</th>
<th>Response category</th>
<th>Response frequency %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Self or independent learner: OUSL student should have the ability to explore knowledge himself. “My palm is only a shade for my head”</td>
<td>36.7% (138)</td>
</tr>
<tr>
<td>2</td>
<td>Prefer conventional role: should be a good group worker, learning collectively, making good relationships among university teacher as well as within students</td>
<td>22.6% (103)</td>
</tr>
<tr>
<td>3</td>
<td>No Idea about student role</td>
<td>19.15% (72)</td>
</tr>
<tr>
<td>4</td>
<td>The OUSL student should be a good time manager: Ability to balance time and manage arise conflicts in different roles (university, working place and home)</td>
<td>6.38% (24)</td>
</tr>
<tr>
<td>5</td>
<td>Exam-oriented mentality: somehow OUSL students should be able to get through exams within a minimum period for future development</td>
<td>5.85% (22)</td>
</tr>
<tr>
<td>6</td>
<td>Not answer the question</td>
<td>3.19% (12)</td>
</tr>
<tr>
<td>7</td>
<td>Should be a hard worker and a target-oriented person</td>
<td>1.32% (6)</td>
</tr>
<tr>
<td>8</td>
<td>Should have developed a balanced personality to solve and tolerate difficulties in the ODL endeavor, should be a different role player (as an employer, as an undergraduate, as a husband/wife/family member)</td>
<td>1.32% (6)</td>
</tr>
</tbody>
</table>

Table IV. BSc new registrants’ perception of “his/her role as an OUSL student”
duration taken for completing the degree. However, result of this study has recognized
that in contrast BSc new registrant of OUSL have sufficient knowledge on ODL concepts
(Table I). Efficient counseling and orientations programs can be recommended for those
who show poor perception of the ODL teaching-learning process.

Generally, the OUSL uses printed material as instruction media and to a lesser extent
electronic aids as supplementary material. More response frequencies (45.7 percent) clearly
highlight the existing condition in relation to the instructional material. The latest statistics
issued by the Telecommunication Regulatory Commission of Sri Lanka show that active
internet users have increased up to more than 5 million and the number of mobile
broadband connections has doubled each year from 2009 to 2015 (Internet usage of
Sri Lanka, 2016). The computer literacy of the younger generation (aged 15-29) in Sri Lanka
is quite high. Of the 18-24 age group 65 percent could manage basic functions of the
computer as opposed to the 25-29 year olds where only 43 percent demonstrated knowledge
in computer usage (De Silva et al., 2013). We found that more than 92 percent of the BSc
undergraduates registered for OUSL are below 27 years age, and as such should be
computer literate. In addition, responses for the open ended part of the questionnaire
recognized the main instruction media of OUSL as “purely internet” (9.73 percent) or a
“mixture of internet and other electronic or printed material” (35.26 percent) (Table II).
The answer may indicate their subconscious preference to use electronic media including
the internet as instructional media. In this context OUSL is still not on par with the use of
technology by the younger generation. Hence, this study recommends the increase of online
teaching—learning processes for BSc undergraduate of OUSL.

Responses to the “role of the teacher at OUSL” can be summarized into six categories
(Table III). The majority of response frequencies correctly recognized the role of an ODL
teacher as “a helper”; a “good counselor,” a “facilitator” or “an adviser” (56.9 percent). But quite
a large frequency percentage (37.5 percent) of responses recognized the role of an ODL teacher
to be similar to the role played by a conventional “teacher” or did not respond to the question.

Most of BSc undergraduates of OUSL are comparatively young and below 27 years of
age. After facing the highly competitive G.C.E. (A/L) examination, most do not gain the
opportunity of entering a conventional state university. From among those who are eligible
for university education, it is actually only 28.87 percent that qualify to enroll at
conventional state universities in biological science discipline and only 38.12 percent in
physical science in the year 2012 (UGC-Sri Lanka, 2016). The majority of students, who
perform relatively well at G.C.E. (A/L) do not get the opportunity to enter state universities,
due to the limitation of resources and state higher education institutes in the country.
Further, 38.05 percent of students, did not qualifying for university/tertiary education based
on G.C.E. (A/L) 2012 results (Department of Examination, 2016). This clearly shows that the
larger percentage of students either do not qualify G.C.E. (A/L) exam or even if qualified do
not get the opportunity to gain (state) university entrance. These students either select
suitable middle class employment or directly enter for further studies at private institutes or
OUSL. Even those who are employed may select higher education at OUSL.

Our study results shows that, 48 percent of new registrants are between the ages 18 and
22; and therefore, fresh school leavers. The role played by secondary school teachers would
still be fresh in the minds of new registrants. Hence they prefer close contact with teachers.
Basically, ODL students are physically isolated. However, psychologically, adolescent
students prefer and need emotional contact with teachers or their peers. Situations such as
these make students distressed or in the long run even give up the study program. It has
been identified that 64 percent of BSc undergraduate of OUSL are psychologically
distressed (Kuruppuarachchi et al., 2012a, b). This could be due to physical isolation or
problems faced during adjustment to the ODL system of OUSL or both. Therefore it is
recommended that appropriate orientation and counseling programs be conducted during
initial registration time (of BSc undergraduates) are strengthened to overcome these issues. The major components of the teaching-learning process of the ODL system are course development and delivery. During course development, more attention is given to develop knowledge or for the cognitive development of the students. The course delivery part of the ODL system therefore, should mainly consider the development of skills, attitudes and values. In the long run the teaching-learning process of OUSL should be reformed to fulfill student’s psychological needs, in addition to the knowledge-based domains. Hence, course delivery components such as practical, day schools, and other extracurricular work should be reformed to provide more interaction between teachers and students as well as among students. For example, peer interaction sessions, active learning, cooperative learning or group learning techniques can be strengthened to develop soft skills and values and enhance social interaction to eliminate the social isolation of adolescent BSc undergraduates of OUSL.

According to the Sri Lanka Qualification Framework (2015), 12 learning outcomes should be achieved by any undergraduate of a Sri Lankan University. This includes only 25 percent for knowledge-based domains and the remaining 75 percent are mainly targeted toward attitude and skills development. Hence, OUSL too has to reform its teaching-learning process to include more attitude and skills development while maintaining its identity as an ODL institute.

The responses on “The role of OUSL students” could be summarized into eight categories. The relevant response frequencies are displayed in Table IV. Results showed that mainly response category number 1 (36.7 percent), 4 (6.38 percent), 7 (3.19 percent) and 8 (1.32 percent) (total 47.59 percent) represent clear ODL concepts on the “role of an ODL student.” However, the responses of students indicate that the majority (52.1 percent) did not have a clear idea of this concept.

Central to Carl Rogers’ personality theory (Rogers, 1959) is the notion of self or self-concept. This is defined as “the organized, consistent set of perceptions and beliefs about oneself.” The self is the humanistic term for who we really are as a person. According to Carl Rogers, the main duty even of a counselor is to help the client understand “himself.” As such the students should be helped to get a more clear perception of their responsibility and role in the ODL system. Students should be made to realize that their success in the ODL setup is clearly dependant on their own efforts, and even more so that in a conventional University setup.

The Sri Lankan school education system is mainly knowledge oriented. The G.C.E. (A/L) science discipline is a highly competitive and exam-oriented teaching-learning process. A “spoon feeding system” has been developed by A/L teachers as well as private tutors. This maybe reflected in the BSc undergraduates’ inability to clearly recognize ones’ role and responsibility as an ODL student. The new teaching-learning process, the physically isolated nature of the ODL system along with other problems faced during adjustment to the new system would seriously affected the new registrant to the OUSL.

ODL students are more likely to have insecurities about learning. New registrant are particularly vulnerable, as they struggle to establish study habits, time management and maintain motivation. The adjustment difficulties, financial difficulties, limited access to resources, and conflict between education and family life or work place, etc., could lead to further insecurities (Owens et al., 2009).

Due to these reasons, the academic and administrative structure of OUSL, especially in the Faculty of Natural Sciences should not be restricted to the “conventional ODL concepts and methodologies.” It should be flexible, to promote students’ interactive sessions, peer learning, group learning, active learning or collaborative learning systems which can be include into course delivery. In addition, the BSc undergraduate should be trained to recognize “self-concepts” and “self-learning concepts” during the time of initial registration by strengthening fruitful orientation and counseling programs. In addition to the limited existing
on-campus orientation and academic counseling programs for new registrants in the Faculty of Natural Sciences, it has recently (2014) introduced “Peer-Assisted Study Sessions”. These are conducted with the help of existing and newly graduated OUSL students as young mentors. This will help students to develop social and academic interactions and share experiences gained by past students (Bandarage et al., 2015). Unfortunately, students’ participation in such academic support work has not been encouraging. Hence, new methodologies need to be developed to improve attitudes and social skills of OUSL, BSc undergraduates by way of introducing group assignments, presentations and case studies, etc. which are considered as a component of the continuous assessments.

Conclusions/recommendations
In conclusion, the majority of BSc undergraduates of OUSL represent the employed, late adolescent lower middle class. Prior knowledge on ODL concepts were developed by the majority of BSc undergraduates while not properly recognizing the “role of an ODL student.” Hence, more attention should be paid to make students aware of their role in the ODL system, in order to succeed at OUSL. In addition, the following recommendation can be highlighted from the findings of the study:

- Increasing online-based teaching-learning processes for BSc undergraduate of OUSL.
- To fulfill student’s psychological needs, the ODL delivery components should be reformed to provide more interaction between teachers and students as well as among students, i.e. active learning, cooperate learning, group learning, case studies, etc. in limited contact sessions.
- A fruitful orientation and counseling program during the initial registration period to adjust to “self-learning,” “ODL methodologies,” “self-management” to work efficiently in physically isolation.
- The teaching-learning system of OUSL should not be restricted to the “Conventional ODL concept and methodologies” and should be flexible to promote possible students interactions.
- The academic and administrative structure of the OUSL especially in the Faculty of Natural Sciences should try to diversify its clientele. While commending the institution for providing opportunities for the less privileged, the university should make an effort to attract clients of diverse social classes for the long-term sustainability of the institute.

The majority of studies on ODL focus on the nature or improvement of the educational process or the output (the graduate). However, as the output depends on both the process and the input (students that enroll) it is equally important to pay adequate attention to this aspect and conduct such surveys on its clientele periodically as well. This will enable the institute to formulate/modify its policies on educational administration, program development, delivery, student welfare, etc. to better cater to its clientele and in the process lead to the betterment of the institution.

References
American Association for Public Opinion Research (2000), Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys, 2nd ed. (electronic version), AAPOR, Ann Arbor, MI.


Corresponding author
K.A.J.M. Kuruppuarachchi can be contacted at: kajmkuruppu@gmail.com

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Analyzing students online learning behavior in blended courses using Moodle

Rosalina Rebucas Estacio and Rodolfo Callanta Raga Jr
College of Computer Studies and Engineering, Jose Rizal University, Mandaluyong City, Philippines

Abstract

Purpose – The purpose of this paper is to describe a proposal for a data-driven investigation aimed at determining whether students’ learning behavior can be extracted and visualized from action logs recorded by Moodle. The paper also tried to show whether there is a correlation between the activity level of students in online environments and their academic performance with respect to final grade.

Design/methodology/approach – The analysis was carried out using log data obtained from various courses dispensed in a university using a Moodle platform. The study also collected demographic profiles of students and compared them with their activity level in order to analyze how these attributes affect students’ level of activity in the online environment.

Findings – This work has shown that data mining algorithm like vector space model can be used to aggregate the action logs of students and quantify it into a single numeric value that can be used to generate visualizations of students’ level of activity. The current investigation indicates that there is a lot of variability in terms of the correlation between these two variables.

Practical implications – The value presented in the study can help instructors monitor course progression and enable them to rapidly identify which students are not performing well and adjust their pedagogical strategies accordingly.

Originality/value – A plan to continue the work by developing a complete dashboard style interface that instructors can use is already underway. More data need to be collected and more advanced processing tools are necessary in order to obtain a better perspective on this issue.

Keywords Moodle, Online behaviour, Activity logs, LMS access, Vector space model

1. Introduction

Many higher educational institutions in the Philippines have started to implement web-based learning environments capable of delivering online education in a blended learning academic setting. Blended learning, also called hybrid learning or mixed method learning involves both face-to-face classroom style instruction as well as the use of online methods (Prasad, 2015). Researchers are unanimous in stating that the blended learning strategy enables educational institutions to implement a more learner-centered approach to teaching where learners are given space and flexibility to indulge with effective learning activities (Alonso et al., 2005; Hughes, 2007; Roby et al., 2013). To implement blended learning, a web-enabled tool or learning management system (LMS) is often utilized to design a particular course in asynchronous mode. Moodle, a free open-source software package used by educators to create online courses (Borromeo, 2013; Maila et al., 2014; Asian Association of Open Universities Journal Vol. 12 No. 1, 2017 pp. 52-68 Emerald Publishing Limited 1858-3431 DOI 10.1108/AAOUJ-01-2017-0016 © Rosalina Rebucas Estacio and Rodolfo Callanta Raga Jr. Published in the Asian Association of Open Universities Journal. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/ licences/by/4.0/legalcode

The authors would like to express their sincerest gratitude to the Office of the Research Director of Jose Rizal University for all the support extended for the accomplishment of this work.
It provides a modular design that makes it easy to add contents that will engage learners and supports a social constructionist pedagogy style of teaching (Romero et al., 2008).

The use of Moodle has been cited by several literatures as an effective tool for teacher course administrative tasks (Perkins and Pfaffman, 2006); improving student inquiry and critical analysis skills (Regueras et al., 2011); inducing self-directed learning (Woltering et al., 2009); as well as promoting collaborative activities (McLuckie et al., 2009). However, even with many cited benefits of using Moodle particularly in higher education institutions, there are still factors that need to be looked upon to ensure its effective implementation. One of the most difficult factors has to do with assessing how the utilization of the various features of Moodle within the online environment affects the overall course performance of students. Are there patterns of utilization that can lead to better success in learning and higher course grade? This aspect can be analyzed by looking into the sort of activities that students often engage with. Due to the nature of the design of Moodle, it is able to routinely collect detailed activity data on students through its log files. Unfortunately, because of the inherent difficulties in handling these enormous log data files generated online by students; teachers would not agree to analyze them manually. Traditional assessment techniques, on the other hand, do not provide appropriate measures on the kind of skills that students develop while interacting with the features of the Moodle environment (Macfadyen and Dawson, 2010).

Fortunately, in the last few years, data mining technologies have been making a lot of headway in capturing and analyzing massive amounts of data (Romero et al., 2008). These technologies utilized techniques adopted from machine learning and text mining which have enabled researcher to gain unique insights from huge amount of data with minimal effort (Blikstein, 2011).

This paper presents part of an on-going research focused on analyzing students’ behavior in a blended learning environment. Hundreds of activity logs for each student were collected, filtered and analyzed using a machine learning technique known as vector space model (VSM). The paper also describes some prototypical coding trajectories generated using these logs, look on probable relationship to student’s overall course performance and finally on effects teaching and learning in blended environments.

2. Previous work
VSM has traditionally been used to search and process important information in large collections of unstructured texts (Raghavan and Wong, 1986; Mikolov et al., 2013; Farid et al., 2016). Recently, however, there has been some progress on utilizing VSM for purposes outside the domain of information retrieval. Sreeja and Mahalakshmi (2016), for example, explored the use of VSM to automatically detect emotions in English poems. They compared the performance of VSM with a probabilistic corpus-based method and found that VSM performs better in recognizing emotions in poems mined from public websites.

Fraser and Hirst (2016) investigated using VSM to detect language impairments among people with Alzheimer’s disease and compare it with those from healthy controls. Initial findings showed changes in word usage in Alzheimer patients after analyzing their words when mapped in the VSM semantic space. Younge and Kuhn (2016) used VSM as a measure to detect patent similarity and concluded that VSM is a better measure to use for this purpose. Li and Zeng (2016) also used VSM as a foundational technology to develop a system that can be used to filter spams in mobile text messages.

The dynamic explosion of information in web-based educational system in recent years has required additional efforts in finding appropriate learning materials suitable for learners. Salehi et al. (2013) developed a hybrid recommender system that can overcome this problem by finding appropriate learning materials based on some specific attributes of each learner.
In the same manner, in this study, we attempt to apply VSM to data generated within web-based environments in the context of blended learning courses to enable instructors to overcome the voluminous amount of activity data generated by students as they interact with resources and with each other within the Moodle system.

Student activity logs are a key resource for gaining insight into student behavior in online environments. Behavior patterns observation, in turn, is a necessary step in detecting students' learning style. Govaerts et al. (2011), for example, developed a tool called student activity meter (SAM) which can visualize the amount of time spent by students on learning activities and resources used in online learning environments. They found that visualizations generated through SAM contribute to creating awareness for teachers and that this awareness enables them to develop various teaching strategies. Ateia and Hamtini (2016), on the other hand, connected students' behavioral patterns to specific features of an online environment then used this to define the effect each visual, auditory, and kinesthetic (VAK) learning style will have on each pattern using the VAK. They claim that web-enabled learning systems "that are supported by a dynamic approach to detect the learners learning styles are better and more effective than traditional ones that extract learner learning style using traditional questionnaires."

Similarly, the work by Romero et al. (2013) mines the web usage data from Moodle to predict the student performance. They use features such as assignment, quizzes, and forum activity to predict students final ratings based on four categories – fail, pass, good, and excellent. The paper also presents a mining tool to extract data from Moodle. The results of the paper compare multiple algorithms and show that the fuzzy rule learning algorithms and decision trees perform well with an accuracy of 65 percent.

Agnihotri et al. (2015) focused on studying login data extensively and used it to cluster the students using the data they generate while interacting with a tool called Connect. The work used machine learning-based clustering techniques to group students based on their attempts, scores, and logins. Their results identified three distinct student clusters: "high-achieving students," "low-achieving students," and "persistent students" (Luik and Mikk, 2008). They also found a non-linear relationship between logins and performance based on the cluster results.

Wen and Rosé (2014) also attempted to look at the varying patterns in the behavior of students relative to their grades. Utilizing clickstream data from MOOC courses to characterize the sessions, they were able to mine student behavior within individual sessions. Results of their experiments show distinctive behaviors among students who pass, fail, and receive a distinction. This provides an indication on how different students with varying levels of course performances distribute their activities differently in online environments.

While many works in the literature describe methods to identify patterns in the student behavior, Champaign et al. (2014) posited that there is a strong negative correlation between student’s skill and the time they spend doing online tasks. They likewise observed a negative correlation between the improvement in skill and the time on task. This finding provided added motivation for the current work in terms of verifying whether similar correlations also exist in and among students exposed in blended courses.

3. Research questions
The main goal of this research is to find effective ways to sift through the vast quantity of data generated by web-based learning environments. In particular, it aims to look into the action log data maintained by Moodle to determine whether processing models can be developed that can extract useful information that instructors can use in monitoring class activity. This involved the extraction of sample log data for selected blended learning
courses offered at Jose Rizal University (JRU). Specifically, the study addresses the following research questions:

**RQ1.** How can students’ learning behavior be extracted and visualized from activity logs recorded by Moodle?

**RQ2.** Can Moodle’s action log of student’s online activity offer meaningful insight into students’ course performance?

**RQ3.** Does the demographic profile of students have any effect on their level of activity in an online learning environment such as Moodle?

### 4. Data set

The data analyzed and processed in this exploratory research was extracted from various blended learning courses offered at JRU during the second semester of SY 2015-2016. “A blended learning course is defined as a formal education program in which a student learns at least in part through online learning, with some element of student control over time, place, path, and/or pace” (Blended Learning Definitions, 2017; Horn, 2013). The courses under study include Elementary Statistics (MAT22), Human Behavior Organization (MGT26), Engineering Management (EGR36), and Ethics in Information Technology (ITC56). These courses are being offered to undergraduate students taking up BSA, BScpE, and BSIT, respectively. These blended courses were chosen because they all served as pilot implementations of the Course Redesign Program (CRP) of JRU where extensive use of the Moodle environment was introduced to enable instructors to deliver course contents – learning materials, supplemental links; promote student engagement thru online forums and chats; assessment tasks – quizzes and assignments more effectively.

As such, the online structure of each course consists of courses – readings, assignments, exercises, lecture quizzes, and a final exam, which the students are required to complete with a minimum grade of 3.5 to pass the course.

In this setup, instructors are still required to spend an one hour lecture time with the students each week, afterward, their tasks consist mainly of monitoring students’ online activities while students, aside from attending class lecture will need to spend a minimum of two hours laboratory time with Moodle every week at their own discretion.

To support this setup, a pre-designed course content template is already placed in the Moodle course resources before the start of each class; however, individual teachers can customize this template by uploading additional learning materials like PowerPoint presentations, video clips, and web resources. They can also require forum participations and add additional exercises, assignments, and quizzes as they deem fit. Students can browse the contents independently through individual accounts. Some student accounts can only be used locally in the laboratories but there are also experimental accounts which are cloud-based and can be used online and accessed conveniently anywhere. Many students even access the online courses using mobile devices.

The predefined course template divides each course into several modules. For each module, students were asked to complete topic-related readings and perform the prescribed exercises, on scheduled dates, and take the online quizzes. Students can also download content materials and exercises and work on them offline.

Relative to this setup, Moodle has built-in features that can produce several types of reports that can be used to track student activity. One of these reports, called action logs, enable instructors to keep track of which resources and activities in a course have been accessed, when, and by which student. For the purposes of this study, logs of students’ action for the entire semester for each of the courses were collected then cleaned up. This resulted in a data set with a total of $n = 199$ students.
4.1 Action logs

Each event record in the raw action log has six attributes (see Table I): course name, time of the event, IP address, username, action, and information. In this study, we only focused on the username and action attributes, the other attributes from the raw data were reserved for future use. The action attribute represents actions initiated by students on various items that can be accessed from Moodle such as assignment, quiz or assessment, course content, forum discussion, resource, and URLs. The actions that can be performed on these items include:

- view individual and view all – opening the items on Moodle;
- view forum – opening the forums;
- forum add discussion – add or post a forum topic;
- submit – upload completed assignments or quiz; and
- submit for grading – submit the uploaded assignments or quiz for grading.

4.2 Extracted log records

Table II provides the total number of action log records extracted for each course as well as the average number of actions per student. These logs constitute actions initiated for Moodle tools identified previously.

4.3 Student demographics

Finally, student demographics were taken by means of a structured survey using purposive sampling, participants in the survey conducted were the same students whose activity logs were extracted and processed. These data as shown in Table III are essential in identifying possible focal determinants of students’ online behavior exhibited by Moodle recorded action logs.

These attributes would be checked later to determine whether students’ online activity is affected by gender, year level, enrollment status or the number of CRP (online) course they are taking for second semester.

The two latter attributes (device ownership and access mode) would help describe how students took advantage of the mobility factor of an online course relative to the blended

<table>
<thead>
<tr>
<th>Data dimension</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course</td>
<td>Identification string of the course in which the action is related</td>
</tr>
<tr>
<td>Time</td>
<td>Date and time stamp of when the action was executed</td>
</tr>
<tr>
<td>IP address</td>
<td>Unique numerical label assigned to the device used by the user</td>
</tr>
<tr>
<td>User full name</td>
<td>The user who initiated the action</td>
</tr>
<tr>
<td>Action</td>
<td>Type of action initiated</td>
</tr>
<tr>
<td>Information</td>
<td>General information on learning activities</td>
</tr>
</tbody>
</table>

**Table I.** Moodle action log data attributes

<table>
<thead>
<tr>
<th>Course</th>
<th>Number of students</th>
<th>Total number of action logs</th>
<th>Average actions per student</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAT22</td>
<td>60</td>
<td>10,370</td>
<td>173</td>
<td>43.76</td>
</tr>
<tr>
<td>EGR36</td>
<td>80</td>
<td>12,009</td>
<td>150</td>
<td>17.04</td>
</tr>
<tr>
<td>MGT26</td>
<td>22</td>
<td>2,687</td>
<td>128</td>
<td>20.32</td>
</tr>
<tr>
<td>ITC56</td>
<td>37</td>
<td>3,520</td>
<td>95</td>
<td>16.61</td>
</tr>
<tr>
<td>Total</td>
<td>199</td>
<td>28,586</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table II.** Total number of actions per course
mode design (Section 4.3). Device ownership, in particular, defines how many students have their own desktop or mobile provisions, while access mode tells us if students are using these for LMS access relative to specific types of Moodle activity, e.g. content-, assessment- or engagement-related tasks.

As shown, the majority of students own a computer with internet access; 58.33, 61.90, 62.50, and 78.38 percent for MAT22, MGT26, EGR36, and ITC56, respectively, and they are those who access Moodle using their own provisions as reflected by 46.67, 76.19, 20.00, and 32.43 percent.

Respondents with no PC or have a computer but no internet access relied on using JRU Open Lab (33.33, 19.05, 25.00, and 18.92 percent). This implies a positive observation because it implies that students who lack personal devices can still access course contents and perform online tasks through the university infrastructure as provided in the open lab.

Use of mobile device was differentiated from use mobile device home/mobile by verifying IP address stamped in each action log (the IP address is a set of numeric values that specifically identifies the device being used by the student). Another notable item in the survey shows that MAT22 students have taken more CRP (Moodle online) courses during the semester compared to the other three courses.

5. Analysis of activity data

5.1 Extracting and visualizing learning behavior

After data collection, the first question that was addressed is how to process the data set and extract patterns of activity that can be used to visualize students learning behavior. Following the mining process described by Romero et al. (2008), as shown in Figure 1, a two-phase process was used which include initial preprocessing of data then afterwards applying data mining algorithms that transform the data into a form suitable for interpretation and evaluation. In the context of this study, Moodle log data were collected from JRU LMS for a particular CRP course as depicted in Figure 1.
5.1.1 Data preprocessing. Data pre-processing is one crucial step in data mining (Mohamed, 2014). In this phase, the raw log files were first processed to clean and prepare it for further processing. This is critical because many of the data sets extracted in Moodle can have missing values, noisy data, and/or irrelevant and redundant information. For this purpose, the raw log files were first imported into an Excel worksheet. Here, the actions logged by instructors and course administrators were selectively removed and the data set was anonymized by removing each student’s name and replacing it with a unique identification number. Processing then started by filtering the data set by course, user identification, and action. Then, two-dimensional tables for each course were built containing the list of student identifiers as row headers and specific types of actions as column headers. Table IV presents the set of action types used in this study for analyzing the students’ online behavior. The key aspect of these actions is that collectively, they can be used to represent the different types of activities that students can engage with inside Moodle, that is: accessing course content, engaging with peers, and taking assessment tests. The key assumption here is that student’s actions indicate intentionality which in turn provide clues, as to their learning preferences. Thus, when categorized based on class activities, the actions helps to infer whether the student prefers to study by accessing learning materials, by engaging with peers and/or the instructor or simply by taking assessment tests.

Each cell in the two-dimensional table was filled with values representing the total number of times each action type was initiated by each student. Figure 2 shows a sample table generated after pre-processing the raw data files. The process of counting this value

<table>
<thead>
<tr>
<th>Class activity</th>
<th>Action type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content access</td>
<td>Course view</td>
</tr>
<tr>
<td></td>
<td>Resource view</td>
</tr>
<tr>
<td></td>
<td>URL view</td>
</tr>
<tr>
<td>Engagement</td>
<td>Forum add disc</td>
</tr>
<tr>
<td></td>
<td>Forum view disc</td>
</tr>
<tr>
<td></td>
<td>Forum view forum</td>
</tr>
<tr>
<td>Assessment</td>
<td>Quiz view</td>
</tr>
<tr>
<td></td>
<td>Quiz attempt</td>
</tr>
<tr>
<td></td>
<td>Assign view</td>
</tr>
<tr>
<td></td>
<td>Assign submit</td>
</tr>
</tbody>
</table>

Table IV. Moodle action identifiers
<table>
<thead>
<tr>
<th>Students</th>
<th>Course view</th>
<th>Resource view</th>
<th>URL view</th>
<th>Post add disc</th>
<th>Forumview disc</th>
<th>Forum view forum</th>
<th>Quiz view</th>
<th>Quiz attempt</th>
<th>Assign view</th>
<th>Assign submit</th>
<th>Final Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1001</td>
<td>16</td>
<td>29</td>
<td>36</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>47</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1002</td>
<td>79</td>
<td>44</td>
<td>7</td>
<td>2</td>
<td>12</td>
<td>10</td>
<td>36</td>
<td>7</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1003</td>
<td>46</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>36</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1004</td>
<td>30</td>
<td>34</td>
<td>8</td>
<td>3</td>
<td>6</td>
<td>13</td>
<td>35</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1005</td>
<td>32</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>1006</td>
<td>20</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>21</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>1007</td>
<td>51</td>
<td>46</td>
<td>28</td>
<td>3</td>
<td>15</td>
<td>23</td>
<td>25</td>
<td>7</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>1008</td>
<td>39</td>
<td>23</td>
<td>15</td>
<td>1</td>
<td>2</td>
<td>7</td>
<td>19</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>1009</td>
<td>93</td>
<td>95</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>28</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>1010</td>
<td>69</td>
<td>18</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>40</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>1011</td>
<td>88</td>
<td>27</td>
<td>9</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>37</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>1012</td>
<td>139</td>
<td>15</td>
<td>21</td>
<td>1</td>
<td>17</td>
<td>13</td>
<td>33</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>1013</td>
<td>63</td>
<td>17</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>8</td>
<td>25</td>
<td>7</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>1014</td>
<td>48</td>
<td>12</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>20</td>
<td>6</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>1015</td>
<td>41</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>39</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>1016</td>
<td>30</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>30</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>1017</td>
<td>35</td>
<td>8</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>24</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>1018</td>
<td>63</td>
<td>35</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>1019</td>
<td>41</td>
<td>25</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>21</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>1020</td>
<td>27</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>23</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Total: 1,150 420 200 13 66 97 69 1.42 0 10 2.743
was automatically done using a customized Excel macro. The total counts extracted for each action type are shown in Table V wherein course views, view forum discussions, view forum, and assignment views have the highest occurrence while quiz view, quiz attempt, add forum discussion, and URL view are relatively low and assignment submit, resource view, and assignment view actions are the fewest actions initiated by the students.

5.1.2 Data mining algorithm – VSM. Data mining algorithms enable extraction and visualization of patterns of activity that can be used to infer students’ behavior.

VSM, a statistical model representation often used in processing documents in information retrieval (Raghavan and Wong, 1986). The main idea behind VSM is to construct vector representation for documents and use these vectors to analyze and compare the contents of each document. A vector is simply a labeled set of values arranged in a specific order. In the case of VSM, the labels are the unique words that occur in the document and the values refer to the number of times each unique word occurred in that document. So for example, if there is $k$ number of documents to be represented and these documents contain $n$ number of unique words, a $k \times n$ matrix can be built as shown in Figure 3. In this matrix, $D_1$ to $D_k$ represent the set of documents while $W_1$ to $W_n$ represent the set of unique words. The values in each cell represent the number of times a specific word $W$ occurred in a particular document $D$. Each row in this matrix is considered a vector representation for its corresponding document.

The analogy used in VSM is that the vector representation acts as a sort of coordinate that can be used to plot the position of the document in an $n$-dimensional semantic space where $n$ corresponds to the number of values in the vector. Figure 4 depicts what a three-dimensional semantic space looks like along with the documents plotted in this space using vector representation.

Using this analogy, to compare the contents of documents VSM simply determines how far the location of their vector representations is within the semantic space. For instance, to determine how closely related the topic of document $D_1$ is to the topic of $D_2$, VSM simply

<table>
<thead>
<tr>
<th>Action</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course view</td>
<td>11,265</td>
</tr>
<tr>
<td>View forum disc</td>
<td>6,491</td>
</tr>
<tr>
<td>View forum</td>
<td>3,047</td>
</tr>
<tr>
<td>Quiz view</td>
<td>2,043</td>
</tr>
<tr>
<td>Quiz attempt</td>
<td>1,870</td>
</tr>
<tr>
<td>Add forum disc</td>
<td>1,817</td>
</tr>
<tr>
<td>URL view</td>
<td>1,193</td>
</tr>
<tr>
<td>Assign submit</td>
<td>444</td>
</tr>
<tr>
<td>Resource view</td>
<td>347</td>
</tr>
<tr>
<td>Assign view</td>
<td>71</td>
</tr>
</tbody>
</table>

Table V. Statistics on logged student actions

<table>
<thead>
<tr>
<th>$w_1$</th>
<th>$w_2$</th>
<th>$w_3$</th>
<th>...</th>
<th>$w_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_1$</td>
<td>14</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$D_2$</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>$D_3$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$D_k$</td>
<td>4</td>
<td>7</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 3. $k \times n$ matrix
measures the distance of \( D_2 \) relative to the position of \( D_1 \). The most common method for measuring this distance is by calculating the cosine of the angle formed between the two locations (represented by the symbol \( \Theta \)). The formula for computing the cosine angle is as follows:

\[
s(x, y) = \frac{x \cdot y}{\| x \| \| y \|} = \frac{\sum_{i=0}^{n-1} x_i y_i}{\sqrt{\sum_{i=0}^{n-1} (x_i)^2} \times \sqrt{\sum_{i=0}^{n-1} (y_i)^2}}
\] (1)

Basically, for two vectors with \( n \) values, this formula simply computes the scalar product of the two vectors for the numerator; computes the product of the length or norm of the two vectors for the denominator. So for example, if vector \( x \) is represented by the values \((1, 1, 1, 3, 0)\) and vector \( y \) is represented by the values \((0, 0, 0, 1, 1)\) the computation for the resulting cosine angle is as follows:

\[
x \cdot y = 1 \times 0 + 1 \times 0 + 1 \times 0 + 3 \times 1 + 0 \times 1 = 3
\]
\[
\| x \| = \sqrt{(1^2 + 1^2 + 1^2 + 3^2 + 0^2)} = \sqrt{12} = 3.4641
\]
\[
\| y \| = \sqrt{(0^2 + 0^2 + 0^2 + 1^2 + 1^2)} = \sqrt{2} = 1.4142
\]
\[
\text{cosine} = 3/(1.4142 \times 3.4641) = 0.61
\]

The cosine formula returns a value between 0 and 1. The rule in VSM is that the more similar the contents of two documents are, the higher their cosine value will be. So a cosine value of 1 for two documents means that the documents are completely identical and a value of 0 means they are totally unrelated. Any value in between reflects the degree of comparison between documents, the higher value means documents are highly related.

5.1.3 Representing student activity using VSM. Given the previous discussion, representing student activity using VSM requires the construction of activity vectors for each student. An activity vector can be defined as simply a list of action types with their corresponding values depicting how many times each action was initiated by the student. Here, a value of 0 means that the action type was not initiated at all. For instance, in Figure 3, the level of activity of student 1001 can be represented by the vector:

| 116 | 29 | 36 | 0 | 0 | 3 | 47 | 8 | 0 | 0 |
There are two ways by which this vector representation can be used. First, it can be used to compare students’ activity to each other in order to group them based on how similar their level of activity is. Second, it can be used to assign students to a predefined set of categories based on how close their activity level vis-à-vis defined activity level for a specific category. Both cases will enable the identification of similar characteristics that occur within each group of students. In this paper, the latter approach is explored.

The color-coded header in Figure 3 indicates the type of class activity to which the action type belongs, such as content access, engagement related, and assessment activity. These sets of activities can be used to classify students to determine which type of activity they implicitly prefer. To do this, an archetypal activity vector for each activity class needs to be constructed. This can be done by setting the corresponding action types for each activity to a non-zero value while the rest of the action type values are set to 0 as shown in Table VI. Thus, the archetypal vector for each activity class would be as follows.

To classify, each student’s activity vector will be compared to the archetypal vector of each activity. The student will be grouped accordingly as per archetypal vector which generated the highest cosine or similarity value. Student activity vector can also be analyzed and grouped and compared on a per course/class basis.

5.2 Analysis of correlation between activity level and course grades
Prior to delving further in analyzing activity logs, it is necessary to first determine whether there is a relationship that exists between the action types initiated by the students and the students’ course achievements, what the direction of the relationship is and its strength of magnitude. For this purpose, the student’s final course grade is treated as an indicator of course achievement; which can reflect both student knowledge and level of engagement. The goal is to gain insight into how students’ actions in the online environment correlate with their course grades. Pearson coefficient correlation ($r$) was used to investigate the significance and computations were done by importing the excel data worksheet to SPSS.

5.3 Analysis of the effect of students demographic profile with level of activity
Correlation and descriptive statistics were conducted to examine whether student demographic attributes, namely, gender, year level, enrollment status, and device ownership could affect the level of LMS utilization as exhibited by total activity logs. Descriptive statistics used to determine mean activity logs of students while Pearson coefficient ($r$) was also used to establish possible relationships.

6. Results and discussions
The development of easily interpretable graphic that can depict trends in student activity based on action logs is a useful tool for instructors to constantly visualize and monitor course progress with minimal effort. Each line point in the graph (Figure 5) shows representative visualizations of the cosine values generated by each student in their respective course. Although students are anonymously depicted, the graph depicts the degree of activity among the participants, and can possibly be even be refined to drill down to each individual student’s level of activity. The visualizations clearly depict some patterns of online behavior relative to three different activities: content access, engagement, and assessment. It indicates that different classes vary widely in how they utilize the tools.

<table>
<thead>
<tr>
<th>Table VI. Vector model for activity class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content</td>
</tr>
<tr>
<td>Engagement</td>
</tr>
<tr>
<td>Assessment</td>
</tr>
</tbody>
</table>
Analyzing students online learning behavior

Figure 5. Visualization of class activity based on action logs analyzed thru VSM
provided within Moodle. And that, even within a certain class, students undertake complex behaviors in allotting time between different tools and activities. Students from MAT22, for example, generally login into Moodle mainly to do assessment tasks with little access to educational resources and engagement. Whereas EGR36 students seem to prioritize content access and engagement over access to assessment tasks; students from MGT26 and ITC56, on the other hand, while more or less showing equal interest on content access and assessment show very little interest in engagement. These visualizations can help course administrators to determine the type of strategic interventions that each course would need to ensure that student’s activities are kept within intended learning outcomes. Unfortunately, while some of these online activities aim to provide effective teaching strategies, the visualizations (along with the cosine ratings) does not seem to correlate with the students’ course accomplishments. This suggestion comes from observing that some course with a low level of online activity (MGT26 e.g. average cosine: 0.33) have higher average grades (e.g. 2.7) than a course with higher levels of activity (e.g. EGR36, average grade: 3.46, average cosine 0.47). This seems to suggest that longer time spent on Moodle may not result in higher course achievement. An implication of these observations is a need to redesign the online component more effectively in order to achieve quality instruction.

Another issue that the visualizations reveal is the lack of a standard teaching approach. Since students’ activities are often governed in part by the teacher requirements, what the visualizations indicate is that teaching methods among different classes seem to vary. Some instructors mainly focus on uploading lectures, while others focus more on assessment tasks or activities; some require a certain level of engagement among their students. More studies are needed to determine which pattern of teaching approach would be most beneficial to the students.

A simple correlation analysis was conducted, in order to determine whether there are relationships between the action types initiated by students and their course accomplishment represented by the final grade given to them by their instructors. As shown in Table VII, the results seem to suggest that there is some variability in terms of positive and significant correlation in the final grade between courses. In Mat22, for instance, the only action which shows correlation with the final grade is the URLView ($r = 0.266$, $p = 0.01$). Whereas, in the EGR36 course, all the actions correlated positively except AssignmentSubmit and AssignmentView ($p = 0.01$). For MGT26, it is the QuizView and QuizAttempt actions that correlated positively ($p = 0.01$) and for ITC56 it is ResourceView and QuizView shows medium to high correlation at $p = 0.01$ and small correlation for AssignView at $p = 0.05$. This variability in terms of correlation, to some degree, seems to agree with our previous observation regarding the lack of correlation between the students’ online activity level and final grades. In other words, in some cases, it may correlate but in others, it may not. The magnitude of correlation also varies from activity type which reflects how students have prioritized tool or activity performed in the online environment. The best explanation for this observation is that instructors are considering other factors in assigning grades to the students which may not be present in the online environment.

<table>
<thead>
<tr>
<th>Course</th>
<th>Res. view</th>
<th>URL view</th>
<th>ForAdd disc</th>
<th>ForView disc</th>
<th>ForView forum</th>
<th>Quiz view</th>
<th>Quiz attempt</th>
<th>Assign submit</th>
<th>Assign view</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAT22 ($n = 60$)</td>
<td>0.245</td>
<td>0.266*</td>
<td>0.183</td>
<td>0.027</td>
<td>0.117</td>
<td>0.145</td>
<td>0.161</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGT26 ($n = 21$)</td>
<td>0.431</td>
<td>0.390</td>
<td>0.258</td>
<td>0.272</td>
<td>0.282</td>
<td>0.742**</td>
<td>0.788**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGR36 ($n = 80$)</td>
<td>0.344**</td>
<td>0.288**</td>
<td>0.520**</td>
<td>0.334**</td>
<td>0.485**</td>
<td>0.415**</td>
<td>0.894**</td>
<td>0.202</td>
<td>0.194</td>
</tr>
<tr>
<td>ITC56 ($n = 37$)</td>
<td>0.604**</td>
<td>0.207</td>
<td>0.045</td>
<td>0.235</td>
<td>0.002</td>
<td>0.470**</td>
<td>0.502</td>
<td>0.373*</td>
<td></td>
</tr>
</tbody>
</table>

Notes: *$p < 0.05$; **$p < 0.01$
The analysis of correlation coefficients for student demographics and total action logs (TAL) obtained as depicted in Table VIII can be observed per course wherein MAT22 has relatively negative, positive small correlation on gender, enrollment status, and device ownership while statistically significant relationship between TAL and CRP courses taken (0.356, p < 0.01, 58). This was contrasted by results for MGT26, with negative/positive medium to high correlation was significantly established for gender and enrollment status (−0.564, p < 0.01, 0.506 at p < 0.05, 19) while EGR36 got a negatively small coefficient for gender (−0.305**, p < 0.01, 78). ITC56 results did not show any significant relationship among any factors considered. Closer analysis of the IP address stamped on the log hits indicated that students still displayed a high level of reliance on computer units provided in the open laboratories. In effect, this nullifies the mobile access advantages of the online course. It was also observed that majority of the task performed online are assessment related. This inference could be attributed to the environment of blended learning implementation of the university where students despite mobility and availability of learning resources online still rely on classroom discussions performed during the face-face session which is one primary structure of a blended mode. Likewise, prior experience in LMS tool indicated by the number of CRP course taken is not a factor on students’ online LMS activity level nor students’ achievement.

7. Limitations and future work

The paper showed how students of various courses utilized the LMS system, it may not be indicative of the overall effectiveness of the system but a structured analytical study of actual online activity thru a data-driven approach gave important highlights. In summary, this work has shown that VSM can be used to aggregate the action logs of students and quantify it into a single numeric value that can be used to generate visualizations of students’ level of activity. While the visualizations do not seem to depict course performance, the value it presents is in terms of helping instructors monitor course progression and enable them to rapidly identify which students are not performing well and adjust their pedagogical strategies accordingly. Since the VSM visualizations generated in this study, for the most part, relied on the structure and contents of the activity reports generated by Moodle, adopting the same methodology to other LMS will necessarily involve minimal adjustments, specifically, in the construction and generation of the activity vectors. Nonetheless, once the necessary adjustments have been made, VSM can be applied to other courses even with varying designs and structure.

A plan to continue the work by developing a complete dashboard style interface that instructors can use is already underway. The study also looked into whether or not various action types can be used as indicators of student’s class performance. The current investigation indicates that there is a lot of variability in terms of the correlation between these two variables. It is highly likely that the design and nature of the course as well as the individual teaching strategy of instructors are introducing other factors that are not present.

<table>
<thead>
<tr>
<th>Course</th>
<th>Mean TAL</th>
<th>SD</th>
<th>Gender</th>
<th>Year level</th>
<th>Enrollment status</th>
<th>No. of CRP course taken</th>
<th>Device ownership profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAT22  (n = 60)</td>
<td>98.72</td>
<td>112.02</td>
<td>−0.032</td>
<td>0.125</td>
<td>−0.230</td>
<td>0.356**</td>
<td>0.085</td>
</tr>
<tr>
<td>MGT26  (n = 21)</td>
<td>73.19</td>
<td>35.81</td>
<td>−0.564**</td>
<td>−0.219</td>
<td>0.506*</td>
<td>−0.138</td>
<td>0.132</td>
</tr>
<tr>
<td>EGR36  (n = 80)</td>
<td>150.10</td>
<td>20.52</td>
<td>−0.173</td>
<td>0.167</td>
<td>0.015</td>
<td>0.191</td>
<td></td>
</tr>
<tr>
<td>ITC56  (n = 37)</td>
<td>54.86</td>
<td>29.28</td>
<td>−0.049</td>
<td>−0.224</td>
<td>−0.125</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** *p < 0.05; **p < 0.01

Table VIII. Correlation analysis of total activity log (TAL) and student demographics
in the online environment. A comparative inquiry on a per subject basis can also be done to explore the effectiveness of a particular course module taken by students in the different disciplinary area. More data need to be collected, and more advanced processing tools are needed in order to obtain a better perspective on this issue.

References


**Corresponding author**
Rosalina Rebucas Estacio can be contacted at: rosalina.estacio@jru.edu

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Quality assurance in learning material development at OUM

Nazrai Ahmad Zabidi, Tai Kwan Woo, P. Rajesh Kumar, Mansor Fadzil and Syarifah Hidayatul Syed Husain
Open University Malaysia, Kuala Lumpur, Malaysia

Abstract

Purpose – The purpose of this paper is to discuss the stages involved in the module development process at the Open University Malaysia (OUM), to make them “print ready,” and how this has evolved over the years with various quality assurance (QA) mechanisms in place.

Design/methodology/approach – The paper builds upon the 12 quality control checks in the various stages of module development from pre-development to actual development and finally to post-development process. This is in line with the QA continuous improvement process of PDCA or Plan, Do, Check and Act.

Findings – Change, through the fine-tuning of QA processes and open door communication, is the key toward achieving quality modules, and hence meeting learners’ expectations.

Research limitations/implications – This paper highlights the experiences of only the Centre for Instructional Design and Technology, OUM in the development of quality learning material for the Open and Distance Learning (ODL) learners.

Originality/value – This paper outlines the step-by-step process in module development, from print to print-ready material, to assist the ODL universities in their mission to provide quality learning material to learners.

Keywords Customer satisfaction, Quality assurance, Learning material

Paper type General review

Introduction

The success and effectiveness of the distance education system “depends heavily on the study material, particularly in the form of printed course material” (Jayaram and Dorababu, 2015). Quality in the Open and Distance Learning (ODL) institutions and their programs are often judged in relation to study material because “student learning is at the center of the ODL experience” (Kirkpatrick, 2005; Hashim, 1999). These are the pivots on which the whole learning pedagogy revolves.

ODL learners in Malaysia are also very concerned about the quality of learning material. This was borne out in a recent survey conducted among post-graduate learners in the Open University Malaysia (OUM) in which “good quality modules or study guides” was listed as one of the five items that were important to the learners (Abdol Latif et al., 2015). The survey concluded that if the course contents, among others, are positively addressed, the learners are likely to continue studying at the OUM and attain their educational goals. At the end of the day, the ODL learners aspire to achieve three main outputs, namely, “gaining new knowledge, acquiring new skills and most importantly obtaining higher qualifications” which they regard as necessary for their social and economic progress (Idrus, 2006).

Historically, quality assurance (QA) first gained prominence in the fields of engineering and manufacturing, as a means of ensuring that quality products are made. The quality gurus have provided a number of definitions of quality, such as conformance to customers’
requirements (Crosby, 1979) and fitness for purpose or use (Juran and Godfrey, 1999). Deming, on the other hand, pointed out that customer’s definition of quality is “the only one that matters” (Deming, 1982). Hence, QA and quality control (QC) are part and parcel of attaining quality. Moreover, QA is “a broad concept that focuses on the entire quality system” and includes quality objectives, whereas QC looks at the micro level and focuses on the process of producing the products (ASQ Statistics Division, 1983). In short, QC can also be viewed as a subset of QA (Eriksson, 2016).

From the perspective of ODL, QA should cover a wide range or areas including curriculum design, content and delivery, and the teaching and learning process (Pitsos and Maila, 2014). In this paper, we will explain and illustrate how the Centre for Instructional Design and Technology (CiDT) at the OUM implements QA in the print module development at various stages. We will also touch on the quality management system (QMS) established in CiDT since it was first awarded the MS ISO 9002:2008 certification in 2005. Finally, we will look into the satisfaction level of learners with OUM’s learning material.

QA in ODL
Around the world, the burgeoning demand for ODL institutions to provide opportunities to learners to pursue higher education has shifted the attention of stakeholders and interested parties to the importance of quality. In light of the increasing competition with conventional universities, the ODL institutions need to offer quality courses and modules that satisfy learners and make them feel there is value for money and the education is a good investment; this in turn ensures that the ODL institutions can continuously “sell and survive” (Koul, 2010) and, at the same time, overall improvements in quality in addition to openness and transparency become the key drivers of “competition and collaboration in education and research” (Anderson, 2008; Ossiannilsson and Creelman, 2012).

With respect to higher education, QA has been defined as the “systematic management and assessment procedures adopted by higher education institutions and systems in order to monitor performance against objectives and to ensure achievement of quality outputs and quality improvements” (Harman, 2000). Alternatively, QA has been defined as “a set of activities or procedures that an organization undertakes to ensure that standards are specified and reached consistently for a product or service” (Robinson, 1995) as well as a set of activities that provide “value for money and transformation capabilities” to ODL students (Gandhe, 2010). In a nutshell, QA “helps disseminate best practices, with the prime objective of leading to overall improvement of higher education systems” (Belawati and Zuhairi, 2007).

In addition, an ODL institution can build its good name by having “a systematic and consistent QA system” resulting in a positive impact on its stakeholders, e.g. more satisfied students, more competent personnel and better public image; meanwhile, students too will benefit as they achieve “enhanced learning outcomes from quality instruction, learning materials and interactions with the institution and its staff” (Kirkpatrick, 2005). Hence, providers need to emphasize quality in various aspects such as products, processes, production and delivery systems (Commonwealth of Learning, 1997). For educational institutions, including the ODL, ISO 9000 assures quality in terms of programs and courses, meeting the intended goals and needs of the users, in particular, the learners (Koul, 2010).

QA related to tertiary education in Malaysia
In general, there are two best-practice QA and accreditation agencies in Malaysia, which are Malaysian Qualifications Agency (MQA) and Scientific and Industrial Research Institute of Malaysia (SIRIM).
MQA was “established under the Malaysian Qualifications Agency Act 2007 by the Ministry of Higher Education” (Malaysian Qualifications Agency, 2009) and is responsible for QA of higher education for both the public and the private sectors in Malaysia. One of the initiatives of MQA is carrying out a rating exercise of Malaysian Higher Education Institutions at the undergraduate level, known as SETARA. In 2011, the ODL category was introduced for the first time to rate the quality of teaching and learning at the ODL institutions in the country, “with 23 criteria and 52 indicators for ODL institutions which include process dimensions such as quality delivery and pedagogy” (Mahsood and Chenicheri, 2016).

MS ISO 9001 is one of the Malaysian Standards developed by SIRIM, whereas ISO 9001 is an international standard that many companies use to ensure that their QA system is established and effective. ISO 9001 demonstrates “an organization’s ability to maintain quality in products and services, meeting the requirements of customers and regulations” (Scientific and Industrial Research Institute of Malaysia, 2016).

QA for the ODL programs and institutions is one of the main concerns among institutions and stakeholders today. This is in line with the Malaysian Education Blueprint (2015-2025) which aspires to see lifelong learning as a way of life for all Malaysians. The blueprint further stipulates that “there will be high quality formal, non-formal, and informal programs in a wide range of disciplines and topics to support both professional and personal development” (Ministry of Education, 2013).

A brief background of the OUM and the CiDT
OUM has had a long history in developing study materials for its distance learners. OUM was the first ODL institution in Malaysia and commenced operations in 2001, beginning with the face-to-face mode of learning aided with printed modules and a learning management system platform. Since its inception, the University has been at the forefront championing an accessible, affordable and flexible approach to learning via the ODL mode. OUM has also been highly ranked in the SETARA ODL category with an excellent Tier 5 rating – a testament to the overall quality of the university (Malaysian Qualifications Agency, 2013).

The university has given working adults in Malaysia the opportunity to pursue higher education anytime and anywhere via the ODL mode. Over the years, the university has grown from strength to strength and this is evident from the fact that presently the OUM has 154,000 cumulative learners, 1,000 international learners and 70,000 graduates. The university also employs about 8,000 part-time tutors, 100 full-time faculty members and 500 support staff. Unlike the other ODL providers in the world, the OUM is unique in that it is owned by a consortium of 11 public universities. These public universities have been generous in lending their prestige and expertise in the various fields; this collaboration has added value to the wide range of programs provided by the OUM. Presently, the University practices a blended pedagogy which comprises self-managed learning, face-to-face tutorials and online learning. This pedagogy is dependent on print-ready self-instructional learning materials, commonly known as “modules.”

At the OUM, the CiDT is primarily responsible for the development of learner-friendly and specially constructed modules (Open University Malaysia, 2008). As of May 2016, the CiDT has developed 1,392 print-ready modules, 366 study guides, 865 hypertext markup language (HTML) modules and 181 video lectures. These HTML-based modules help to enrich the learning experience with colorful graphics, interactive activities and exercises, embedded audio and video files as well as links to other online learning resources. Since the HTML modules and video lectures are sourced from printed modules, it is vital for the module development process to undergo a stringent QA process.
On October 2005, the CiDT was among the four learner support service centers in the OUM to be awarded the prestigious MS ISO 9001:2000 certification (Ali and Fadzil, 2013). The growing number of learning materials made it vital for the center to embrace a QA culture. Moreover, the center is the largest unit in the university comprising more than 60 staff with expertise from various disciplines.

The importance of quality print modules in the OUM

When the CiDT opted to pursue the MS ISO 9001:2000 certification, the scope was limited to printed learning materials for programs offered by the OUM. Currently, there are five core activity procedures for print modules:

1. acceptance, selection and appointment of subject matter experts (SMEs);
2. instructional and graphic design;
3. desktop publishing;
4. management of module writing and moderating workshop; and
5. measurement of customer satisfaction.

Ideally, quality learning materials in an ODL setting must be “simple to interact with and should be understood by learners,” so much so that the materials should be self-explanatory, requiring minimal or even no assistance from tutors (Adegoke and Oni, 2015). It has been shown that learners who scored poorly on reading tests were more likely to drop out before completing their courses (Macdonald-Ross and Scott, 1996). If learners are unable to understand the module, they may underperform in their exams, lag behind in their studies, become demoralized and eventually contribute to the high attrition rate of learners at the ODL institution. In fact, the factors for the high attrition rate among ODL learners include “study material in tough or difficult language and delay or non-receipt of study material” (Kumari, 2012). Meeting deadlines is also very important, as failure to achieve the set deadlines can affect other subsequent processes and lead to “devastating consequences” (Belawati and Zuhairi, 2007).

In addition to this, QA is very much output-outcome driven. Initially, the OUM identified three outputs of QA in the ODL for the University, namely, “graduates with new knowledge, graduates with new skills and graduates with new qualifications” (Idrus, 2006). New skills here include language skills in the English language and relevant skills, such as skills in Finance that will allow learners to perform finance-related calculations. As the learners are mostly working adults, the outcomes among others include job promotion, respect in the community and the pursuit of new interest. Additional relevant skills acquired after graduating from the OUM will be part of the additional list of accomplishments to show value that will be useful for “moving up the corporate ladder” (Zhang, 2014).

Hence, the accuracy of the content in learning material is crucial. Information must be up-to-date and new changes must be taken into consideration as these will “ensure the appropriateness and quality of the learning materials” (De Fazio et al., 2012). Therefore, there should be zero-tolerance for errors such as grammatical error, factual error or mistakes in formula or calculations in the learning material. Learning material which is difficult to understand will contribute to high attrition (Kumari, 2012), whilst “inaccurate and outdated learning material” will not make the learners achieve the desired outcomes (Idrus, 2006).

All five of the SOPs were designed to ensure that QCs are in place so that the print-ready modules are of benchmark standards that satisfy the learners’ expectations. Since the CiDT was first awarded the MS ISO 9001:2000 certification in 2005, the center has undergone a number of quality audits which saw corrective and preventive actions (CAPAs) implemented by way of PDCA or Plan, Do, Check and Act. This requires mastering each
process, knowing how to govern it in a way that enhances communication, transparency and accountability. Understanding and analyzing the root cause of problems is imperative to prevent the same problems from recurring (Olzak, 2008). This calls for stringent QC measures which will be discussed in the next section.

The QA process for print module development
At the CiDT, there are three stages involved in the QA of print module development. These include the pre-development, actual development and post-development stages. Figure 1 shows all the stages and the respective QC measures that will be explained in the following subsections.

QA at pre-development stage
In this stage, there are four QC measures in place to ensure that the OUM modules are of benchmark quality.

The first QA check at this stage is the selection of module writers. This has always been stringently done. The modules are written by writers who are (SMEs) nominated by faculty and vetted by the CiDT to ensure that they have the right academic credentials and relevant work experience in their respective fields. Following ISO requirements, the records of all writers are well-maintained and documented.

The second QA check is the module writing and moderating workshop. Upon appointment by the CiDT, the writers and moderators are required to attend a module writing and moderating workshop where they are briefed on general principles of instructional design, house style and editorial format, timelines and expectations, graphics, referencing and citation, as well as the rubrics for evaluation and payment. Attendance at this workshop is a pre-requisite for module development at the OUM, as everything begins with the writers and moderators who must be clear what the requirements and expectations are. Furthermore, writing modules for ODL learners is “an art in itself” because its interactive styles and techniques are distinct from conventional textbook writing (Rahman, 2006; Hashim, 1999).

At this workshop, the writers and moderators are also introduced to the module development team from both faculty and the CiDT so that they feel a sense of collegiality and esprit de corps for the journey that lies ahead. The development of learning materials by a team of SMEs, academic moderators, module writers, tutors and consultants is part of the

<table>
<thead>
<tr>
<th>Pre-Development Stage</th>
<th>Development Stage</th>
<th>Post-Development Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Stringent selection of Module Writers</td>
<td>- Instructional Design</td>
<td></td>
</tr>
<tr>
<td>- Module Writing and Moderating Workshop</td>
<td>- Editing</td>
<td></td>
</tr>
<tr>
<td>- Screening</td>
<td>- Graphic Design</td>
<td></td>
</tr>
<tr>
<td>- Moderated by an SME</td>
<td>- DTP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Internal Review</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Faculty Review</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- QC Review</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Writer's Review</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1.
QA process for print module development in CiDT
The OUM strategy of giving priority to learners and providing them with quality services and products (Kirkpatrick, 2005).

The third quality check is enforced once the module content comes in. This comes in the form of a plagiarism/similarity check via Turnitin© software to ascertain whether the content submitted by the writer is original and not taken from other sources. Turnitin© is a popular web-based plagiarism detection software used in Malaysia and abroad that is “able to match the work submitted by a module writer against works done by other writers from its huge database repository” (Rogers et al., 2009). If the content is found to be highly plagiarized, it will then be returned to the writer for amendment.

Besides plagiarism, the raw content will undergo a screening process conducted by the instructional designers (IDs). Detecting plagiarism is only one aspect of screening. Furthermore, the IDs will ensure that learning outcomes incorporate the Blooms Taxonomy (a progression from lower order to higher order thinking skills) and the desired outcomes. This allows the learning outcomes to be stated in “measurable behavioral or operational verbs” for the learner to monitor his or her own progress (Koul, 2010). Behavioral verbs have two distinct advantages. First, we can use them to “devise activities that make the student an active participant in his or her learning” and second, “the verbs may be used to help to translate the course goals and objectives into situations that are more concrete to the learners and others” (University of North Carolina at Charlotte, 2016). At this stage, contents are checked against the table of contents (TOC) to ensure that they are in accordance with the TOC given by the faculties. References are standardized according to the APA style used in the OUM. A ten-item checklist identifying the elements that need to be present in the content is used as a tool by the IDs throughout the screening procedure. If there are any issues such as plagiarism, improper learning outcomes or improper in-text citation and references, the content will be immediately sent back to the writer for amendments.

Only after these three QC checks have been carried out is the content forwarded to a moderator, an SME in the same field who is nominated by faculty and appointed by the CiDT. The role of the moderator is to ensure that what has been written follows the specifications spelled out by faculty. So, in this sense, the content is vetted for breadth and depth in terms of scope and coverage by a second expert in the field. Moderators are selected based on specific criteria, such as being of similar, or higher, academic standing as the writer. Professionalism is a necessary ingredient in the production of quality learning material (UT, 2002); hence, the proper selection of competent writers and moderators is taken seriously at the CiDT. If the moderator feels the content is lacking or inappropriate in any way, the writer is obliged to make the necessary amendments based on the moderator’s feedback.

**QA at the actual development stage**

This stage sees the involvement of IDs, editors, graphic designers (GDs) and desktop publishers (DTPs). At the CiDT, there are currently 11 IDs, 5 editors, 12 GDs and 10 DTPs who have expertise in different disciplines. In addition to these different units, there are four dedicated Project Coordinators or PCs for each faculty at the CiDT, and they are responsible for liaising with both internal and external parties in module development for their respective faculties and in coordinating the movement of modules to completion. The PCs play a critical role in the overall coordination of module development as they track the progress of each and every module as early as the pre-development stage, to ensure the modules are delivered within the stipulated timelines given to the writers in the appointment letter. The whole module development process can be likened to “the procedures in an assembly line” and any delay at the preliminary stage can have a devastating outcome on subsequent downstream stages (Belawati and Zuhairi, 2007).
At this stage, the module content is first assigned to an ID who will develop the module based on the quad alignment model shown in Figure 2. All the four components of the model, namely, body of content, learning outcomes, assessment tasks and learning activities are explained and illustrated during the module writing and moderating workshop. Besides understanding each component, it is vital for all the components to be aligned with each other, that is “support and be directly relevant to each other,” thereby helping the learners to easily understand the modules (Fink, 2007).

Hence, the IDs and the writer have to work closely to ensure that every topic is “quad-aligned” so that the learners will find the module easy to comprehend. Apart from writers, the relevant faculty members who are known as program coordinators also provide academic input to beef up the content and ensure that the learning outcomes in each topic are elaborated adequately.

The content in a module must be designed in a systematic way to facilitate learning without regular supervision of a tutor. The content must also be interactive with “an element of guided didactic conversation that can create warmth, closeness and two-way communication between learner and materials or between learner and the module writer” (Holmberg, 1977). Toward this end, IDs have specific roles to play in ensuring that “the learning outcomes are clear to the learner, content is arranged in the proper sequence, there are adequate activities, diagrams and illustrations to aid learners and also assignments to regulate learners’ progress” (Koul, 2010). In some cases, the ID may suggest to the writer parts of the content that should be revised or rewritten to better meet the learning needs of learners. For instance, a well-written introduction to a topic can grab a learner’s attention and arouse his/her interest to continue reading the module.

Among the activities that can be found in the OUM modules are self-checks and self-tests, i.e. self-assessments; for post-graduate learners, the modules include several case studies as well. These are designed to make learning interesting, interactive, meaningful and collaborative. Some of the activities also encourage learners to communicate with their peers through the e-forum and post their answers on myINSPIRE, the learning management system of the OUM.

The content is then edited to improve the language, flow and readability of the content without compromising on the original meaning. The upshot of this is that the content becomes more readable, self-explanatory and easier for the learner to understand (i.e. written in a user-friendly language), requiring minimal tutor guidance. The editor also ensures that the content meets the in-house style standards and verifies headings, statistics and data in graphs and figures. After the editing process, the IDs then work closely with GDs to visualize and incorporate graphic elements to enhance the overall visual presentation of the content. A picture is said to be worth a thousand words; thus, the GDs play an important role to ensure the graphics and diagrams can make the module content more coherent to the learners. Finally, the content is sent to the DTPs for layout in the OUM in-house style to maintain the OUM identity. The DTPs also combine text,

**Figure 2.**
The quad alignment model

QA at the post-development stage

Once the actual development stage has been completed, the ninth QC check kicks in as a hard copy of the module is printed out and sent to the ID for a process called internal review (IR), to ensure that everything has been put together correctly. Here, the IDs will perform a final review of the content to verify that the learning outcomes, content, learning activities and assessments are in sync with each other and to make sure that the module is error free.

After amendments based on the IR have been done, the module is sent to faculty for review. At this stage, the program coordinator or the vice-dean of learning materials will vet the module prior to sign-off for print or upload. Amendments may still be requested at this stage.

Once a module has been signed off, a dedicated QC team at the CiDT vets through the final copy once more before it is printed or uploaded. This QC team comprises the director, academics and senior staff at the CiDT who have strong language skills and are well-versed in module development. To standardize the QC process, the CiDT has come up with a checklist to be used by the team. Some of the elements listed in the checklist include checking and correcting the following:

1. the credit page for the latest version;
2. language errors;
3. figures, tables, activities and self-check questions in terms of suitability and relevance; and
4. alignment and indentation.

After this stage, a soft copy of the final module is sent back to the writer for feedback, on the understanding that if there are no amendments from the writer within a week, the module is deemed approved and ready for print or upload.

According to the quality objectives of the CiDT, the process development must be completed within four months of receiving the content from the moderator. As mentioned earlier, deadlines must be met in order to prevent delay of subsequent activities such as timely delivery of modules to learners.

From print to E

With effect from May 2016, the OUM no longer produces print modules. This is in line with global trends and initiatives to completely migrate the learning content toward E. Instead, e-modules are made available on the learning management system of the OUM known as myINSPIRE for learners who wish to download them in PDF format. Learners can either download the entire module or download each topic separately, which has been made possible after the process of chunking. Meanwhile, bookmarking allows learners to navigate to a particular topic or subtopic within the module by clicking on the TOC.

There is also a module feedback mechanism in place on myINSPIRE, whereby learners can convey suggestions to the CiDT – this feedback is then channeled back to the faculty, who may request for module enhancement, in line with the PDCA model.
A module which is deemed stable is categorized as Red Spine. As a general rule of thumb, a Red Spine module is used for about three years before it is again looked into, in keeping with the philosophy of continuous improvement. Knowledge is the “currency a university deals in,” and this knowledge is constantly upgraded to remain current and abreast of changes (De Fazio et al., 2012). The role of the CiDT is to assist the university in this task.

The entire module development process is presented in Figure 3.

Customer feedback
The commitment from the OUM top management throughout the years has been a critical factor in driving the quality culture at the CiDT and the organization as a whole. In line with this commitment to quality, the CiDT conducts a customer satisfaction survey at the end of every academic cycle to gauge feedback regarding the learning material developed and used in that cycle. The benchmark was set at 80 percent by the CiDT management.

Consequently, as shown in Table I, customer satisfaction surveys have shown the satisfaction results to consistently hover above the benchmark of 80 percent for the past four years. Meeting or exceeding 80 percent of satisfaction for each production cycle is another quality objective of the CiDT.

Quality learning can be said to be achieved when the feedback indicates that the outputs and outcomes of the modules are satisfactory (Idrus, 2006). Even if the benchmark set is not at 100 percent, it does not mean that there is a lack of quality as “quality assurance does not expect perfection”; rather there is an expectation from the CiDT to commit toward improvement (Pitsoe and Maila, 2014). The feedback gained from the survey as well as other feedback mechanisms will provide useful input toward PDCA and further enhance the print module development process.
Conclusion

The increase in customer satisfaction shown in Table I demonstrates the positive impact that the QA culture has on the print module development process at the OUM. CiDT has been adopting a stringent auditing process yearly, since 2005, and the continuous improvement efforts have produced a desirable result. These audits include self-audit, internal audit and, finally, quality audit by SIRIM. The QA process for print module development is very clear in defining what the standards of achievement are, including documented procedures for each process.

The QMS in the CiDT is crucial to ensure the continuity of the quality work culture. The various procedures and work instructions that are in place set out a number of tools to facilitate the module development process. Checklists are available in most processes such as screening, conducting module writing and moderating workshops and also during the actual development stage. If followed correctly, the staff will not leave out any important work instruction and the processes can be executed smoothly and effectively.

Besides checklists, the CiDT has a number of feedback processes in place. Clause 8 of the MS ISO 9001:2015 requires the QMS to establish proper measurement, analysis and improvement (Department of Standards Malaysia, 2016). Implementing quality programs such as ISO certification encourage higher education institutions in Malaysia to collect data, which will “enable them to measure progress in key areas and establish benchmarks” (Sohail et al., 2003). The screening checklist for instance also doubles up as a feedback mechanism as it allows IDs to evaluate whether the writers fulfill the requirement set by the CiDT. After the module writing and moderating workshop, the participants consisting of module writers and moderators are given a questionnaire to gauge their level of satisfaction and also to gain input from the participants regarding the workshop. The customer satisfaction survey shown earlier in Table I is also important to gain vital feedback from the ultimate customer – the learners. In addition to these, there are frequent meetings with the faculties to discuss and thrash out issues pertaining to progress in developing the learning materials. Apart from learners being satisfied with the modules, the respective faculty or in this case program coordinator from the faculty, should also be satisfied with the outcome of the modules (Idrus, 2006).

As a result, many potentially serious issues that can lead to non-conformance have been detected and successfully resolved through CAPA. The initiatives introduced after the ISO certification include the screening process, the use of Turnitin software and the setting up of a QC team. Several tools were added to the processes from the pre-ISO stage, such as providing a step-by-step guide to participants for the module writing and moderating workshop and detailed work instructions for all staff. These have led to continuous improvement by reducing wastage in terms of redundancy and having a clearer workflow.

<table>
<thead>
<tr>
<th>Semester</th>
<th>Customer satisfaction (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>September 2012</td>
<td>73.5</td>
</tr>
<tr>
<td>January 2013</td>
<td>82.8</td>
</tr>
<tr>
<td>May 2013</td>
<td>83.5</td>
</tr>
<tr>
<td>September 2013</td>
<td>83.9</td>
</tr>
<tr>
<td>January 2014</td>
<td>81.8</td>
</tr>
<tr>
<td>May 2014</td>
<td>80.7</td>
</tr>
<tr>
<td>September 2014</td>
<td>86.0</td>
</tr>
<tr>
<td>January 2015</td>
<td>87.1</td>
</tr>
<tr>
<td>May 2015</td>
<td>89.55</td>
</tr>
<tr>
<td>September 2015</td>
<td>71.80</td>
</tr>
<tr>
<td>January 2016</td>
<td>83.2</td>
</tr>
</tbody>
</table>

Note: Module feedback form from September 2012 to January 2016 semester cycle
The road ahead is full of challenges as the OUM migrates to more modules that are fully online because the modules must be of good quality to meet learners’ needs and expectations. Migration also indicates that new QC measures must be established to ensure that the fully online modules meet or exceed the learners’ expectations. We further plan to study the impact that the migration to fully online modules has had on the customer’s satisfaction index, and data are currently being gathered for this purpose. The risks are high as the global education platform becomes increasingly competitive and learners more discerning. But at the same time, opportunities abound for the CiDT to assist the University in its mission to provide quality learning materials and make education accessible to all.

References


Holmberg, B. (1977), *Distance Education: A Survey and Bibliography*, Kogan Page, London.


UT (2002), Sistem Jaminan Kualitas Universitas Terbuka (Quality Assurance System of Universitas Terbuka),Universitas Terbuka, Jakarta.


**Corresponding author**

Nazrai Ahmad Zabidi can be contacted at: nazrai@oum.edu.my

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com
The online examinations at Universitas Terbuka: an innovation diffusion viewpoint

Sugilar Sugilar
*Department of Mathematics and Natural Sciences Education, Universitas Terbuka, Tangerang Selatan, Indonesia*

**Abstract**

**Purpose** – The purpose of this paper is to provide insight into the students’ adoption rate of the online examinations at Universitas Terbuka (UT), which is a hotbed of technological innovation in higher education, by analyzing the students’ adoption rate of the online examinations related to their academic factors (program of study, GPA, credit earned), socioeconomic status (age, sex, marital status, employment status), and residential factor.

**Design/methodology/approach** – This is an in-depth case study of 1,540 first-year students for the period ranging from the start of the second semester of 2013 to the first semester of 2016 at Bengkulu Regional Office of UT. During the term, quantitative data were set up with ordinal regression to measure the extent to which the adoption rate categories were influenced by such demographic characteristics as academic attainment, personal background, and current residence.

**Findings** – The results showed that while the program of study, grade point average, credits earned, sex, and residence had systematic effects on adopter categories, employment and marital status did not. The highlights of the results were that students with GPA less than or equal to 2.9 were 2.02 times likely to fall into a higher adopter category as compared to those with GPA above 2.9, and students residing in city were 2.50 times more likely to hit the higher levels of adopter category, compared to those residing outside city.

**Research limitations/implications** – As this is a case study of the students enrolled in the second semester of 2013 and because there has been a rapid change in the way people access information technology, further work should be done, in particular on the sample of students who have enrolled lately.

**Practical implications** – The institution of open and distance learning (ODL) should accelerate the introduction of new learning resources based on diffusion of innovation modeling.

**Social implications** – It is also recommended to especially encourage the new students of ODL to have a feeling of easiness and self-confidence regarding online examinations, and understand their importance. To achieve this objective, regional office staffs can guide the students to try out the online examination in the orientation study activities for new students.

**Originality/value** – Despite the extensive research on diffusion model for the past decades, this field of study has much more to offer in terms of describing and incorporating such current innovation as an online examination in ODL platforms with which to associate students’ academic and demographic profiles.

**Keywords** Diffusion of innovation, Adopter categories, Adoption rate, Online examination, Open and distance learning university

**Paper type** Research paper

**Introduction**

Online examination is a relatively more open and flexible learning platform, and is more convenient than a pencil-and-paper examination. As such, open and distance learning (ODL) institutions are urged to lead the endeavor to provide quality and excellence in the online examination services. Khare and Lam (2008) argued that “the implementation of online examination for distant learners can no longer be ignored or indefinitely postponed” (para. 1).
Conforming to such stipulation, Universitas Terbuka (UT) conducted the online examination in 2010. The number of participants saw a considerable rise from 332 in the first semester of 2010 to a staggering 2,293 in the second semester of 2012 (Sapriati and Pardeke, 2015). This paper reflects on the proliferation of online examination as a newly introduced education platform and analyzes how and why it broadened the extent to which the aforesaid students’ demographics were interrelated, to bring those demographics into a unified diffusion framework (Iftakhar, 2016).

Lane and van Dorp (2011) defined the diffusion of innovation as a process of adoption involving four elements, i.e. an object of innovation, a group of people, communication channels, and a period of time. Accordingly, the diffusion of innovation is an adoption process of an innovation object in a group of people through communication channels over a period of time. This paper weighs in the concept of innovation-decision process that the students of UT depend upon to make decisions about rejecting, suspending, or using and adopting the online examination as an innovation in UT. Rogers (2003) has broken the innovation-decision process into five stages:

1. Knowledge – the individual (or other decision-making unit) is exposed to the innovation existence and gains some understanding of how it functions; (2) persuasion – the individual (or other decision-making unit) forms a favorable or unfavorable attitude toward the innovation; (3) decision – the individual (or other decision-making unit) engages in activities that lead to a choice to adopt or reject the innovation; (4) implementation – the individual (or other decision-making unit) puts an innovation into use; and (5) confirmation – the individual (or other decision-making unit) seeks reinforcement for an innovation-decision already made, but he or she may reverse this decision if exposed to conflicting messages about the innovation (p. 206).

The innovation-decision period is referred to as adoption rate, which in this paper attributes to the students’ time elapsing from awareness-knowledge of the online examination to the point at which it is put into use.

ODL has gained tremendous growth with an ample amount of innovations, most notably, in the strict sense, for students’ learning support. UT as an ODL institution should accordingly keep providing sustainable innovations of learning support to its 300,000 students currently studying in nearly 13,000 islands in Indonesia. There has therefore been an urgent need for ODL institutions to observe and integrate the online examination as a modeling effort of diverse types of interpersonal characteristics, which are the students’ demographics in this case, into a valid single framework of diffusion. Research in the diffusion modeling will have to expand the horizon. This paper proposes potential guidances for that expansion.

The purpose of the study was to elucidate how the adoption rate of online examination as an innovation at Bengkulu Regional Office of UT was linked to a number of students’ demographics. The length of time between the beginning of the second semester of 2013 and their first participation in the online examination was taken into account to extend the measure of the students’ adoption. The diffusion of innovation theory was then used to analyze the students’ classification based on the adoption rates, i.e. innovators, early adopters, early majority, late majority, and laggards.

On the basis of the aforementioned purpose of study, the researcher therefore outlines research question as follows:

RQ1. How is the students’ adoption rate of online examination related to their academic factors (program of study, GPA, credit earned), socioeconomic status (age, sex, marital status, employment status), and residential factor?

Research context

UT has provided two forms of students’ learning outcome evaluation, known as the final evaluation, i.e. paper-and-pencil test and online examination. The paper-and-pencil test was organized simultaneously in a fixed schedule and place determined by UT as written in the
academic calendar. The online examination was intended for the students who wanted to take the final evaluation beyond the schedule of the academic calendar of UT. In UT’s regulation, the online examination is an option for the students to partake the final evaluation. Bengkulu Regional Office of UT began administering the online examination since the second semester of 2013. In the first semester of 2015, the number of students participating in the online examination was 1,239 or about 30 percent of the total number of students in the educational program for non-teachers. Since the first semester of 2016, re-examination has been introduced as the online examination in the event of students failing paper-and-pencil examination. In such capacity, the number of the participants is most likely to continue to have a growing number in the future.

The data showed that while the online examination was still carried out at Bengkulu Regional Office of UT in Bengkulu City, the capital city of Bengkulu Province, 66 percent of the participants came from outside Bengkulu City. They consequently had to make a travel arrangements, as they stayed hours away from the exam venues, so as to manage to snag a spot. Students from Mukomuko Regency, as an example, had to go on a seven-hour trip to get to Bengkulu City. A research result found that the cost factor of commuting to the exam venues was a determinant factor, though it might depend on the commuting distance and proximity to the exam venues (Sugilar, 2016). Despite the travel and financial burden, most of the students had high academic aspiration and a strong commitment to the university.

Prior to sitting for an online examination, students have to meet a number of procedures. Using their logins, students have to register for the desired examinations. It is their responsibility to ensure that they are aware of the examination details, including the date, the starting time, and the venue, set out by Bengkulu Regional Office of UT. Hence, it is highly recommended that they contact the regional office to get notified at all times about the online examination arrangements.

Related studies
Socioeconomic status has effects on the adoption rate. A study by Ganiyu et al. (2013) described that socioeconomic characteristics of adopters were closely associated with the adoption of innovation, which in turn heavily modulated it. Zhang et al. (2015) studied the adoption process of the e-appointment service focused on some factors influencing patient acceptance and the usage of this process, and found that “the patients’ social and demographic characteristics, including age, education level, and work status, appeared to have influenced their choice of use or non-use of the e-appointment service” (p. 13). The research also found that the variable of sex play-acted as a moderating variable on the correlation between employment status and usage of the e-appointment service. Accordingly, a recent study by Penjor (2016) used age and gender to predict the adoption rate for the virtual learning environment among academic staffs. To conclude, in some cases of adoption process of an innovation, the socioeconomic status variable was related to the level of innovativeness of the adoption. This paper assumed that the socioeconomic status had an effect on the adoption of the online examination and conceptualized socioeconomic status which included the variables of age, sex, marital status, and working status.

Several personal characteristics have explained the individual adoption rate of an innovation. Rogers (2003) generalized that “earlier adopters had greater intelligence, […] more favorable attitude toward education and science, […] had higher levels of achievement motivation than later adopters, [and] […] had higher aspirations (for education, occupations, and so on)” (p. 258). A study by Campbell (2015) about the diffusion of innovation of video-conference technology in distance education pointed out that the interaction of the students with instructors and material were key elements affecting adoption decisions of students regarding video-conference. In an academic setting, as discussed in this paper, students’ pursuance of a program of study, GPA, and credits earned may attribute to such characters.
The students' program of study, among other variables in this paper, was included due to an assumption that the characteristics of students of teacher-training programs were different from those of the regular programs. The students of teacher-training programs are typically professional teachers who are required by the law to earn a bachelor degree, while those of regular programs are non-teachers, be it employees or fresh graduates from high school, who attend UT with considerably diverse motivations and goals which they might want to achieve.

Another characteristic equally relevant to this study is the degree of cosmopoliteness, i.e. "the degree to which an individual is oriented outside the social system" (Rogers, 2003, p. 259). Rogers specified that early adopters are more cosmopolitan than later adopters, since the innovators' networks are more likely to be outside, rather than within, their social system and "they travel widely and are involved in matters beyond the boundaries of their local system" (Rogers, 2003, p. 258). On the basis of this theory, this study broadened in scope from focusing on socioeconomic status to encompass the so-called cosmopoliteness with respect to the variable of residence, which denoted whether the respondents currently resided in or outside Bengkulu City.

Despite the growing trend of the innovation, online examinations at UT have yet to garner much research attention to date. One of the researches was regarding a number of determinants of students’ participation in the online examination by Sugilar (2016), who documented group comparison analysis between the groups of participating and non-participating students. The results showed that the following factors influenced the students to participate in the online examination: self-efficacy in using computers, perceived easiness in operating an online examination, perceived importance of online examination, intrinsic value of online examination, and cost of online examination. In addition, the following personal factors of students were of equally crucial importance in relation to their participation in the online examination: age, grade point average, gender, and marital status. Based on the rank of the amount of effect size for each factor, the order of the five most important of those nine factors was as follows: first, perceived easiness, second, intrinsic value, third, self-confidence, fourth, marital status, and fifth, grade point average.

Another research was conducted by Iriani (2010) who evaluated the online examination practices at Surabaya Regional Offices of UT. The results concluded that 64.22 percent of the students had not known the online examination, and only 2 percent of the students thus 2 percent participated in the online examination. However, the students who participated in the online examination showed exceptionally favorable attitudes toward the arrangement of the online examination.

Sapriati and Zuhairi (2010) observed that in a regional office of UT about 75 percent of students who took a computer-based testing were also participating in a paper-and-pencil test. For the students taking both examination modes, Sapriati and Zuhairi found that the mean score of computer-based testing (equal to 54.00) was higher than the mean score of paper-and-pencil testing (equal to 46.00). Since, as the students know certainly, the examination results do not rely on the mode of examination, rather they depend on the learning process before the examination, these findings expressed that the students had a good preparation before they took the computer-based testing; this was probably because the paper-and-pencil testing was held prior to the computer-based testing.

Research method
This case analysis of 1,540 first-year students at the beginning of the second semester of 2013 at Bengkulu Regional Office of UT employed ordinal regression (Norusis, n.d.) to examine the adoption rate among students of the online examination based on explanatory factors and variables, i.e. academic factors (program of study, GPA, credits earned), personal factors (age, sex, marital status, working status), and residential factor (within or outside the city). The adoption rate of the online examination as an innovation was measured for each student.
The adoption rate was measured by the number of semesters between the beginning of the second semester of 2013 and the first time they participated in the online examination until the first semester of 2016. Because the number of semesters from the first semester of 2013 to the first semester of 2016 was seven, the adoption rate was represented by \(7 - d\), where \(d\) was the number of semesters between the beginning of the second semester and their first participation in the online examination. On the basis of this formula, it can be identified that students who were first participating in the online examination in the second semester of 2014 had the adoption rate equal to 7 – 2 = 5, since the students needed two semesters before they took the online examination. The students who participated at the time of their first enrollment (the second semester of 2013) would have the adoption rate equal to 7 – 0 = 7. In this case, the students who never participated in the online examination, until the first semester of 2016, were considered to have the adoption rate equal to 7 – 6 = 1. Therefore, there were seven adoption rate categories, from 1 to 7. Whenever the students showed higher adoption rate, they were more likely to adopt the online examination faster.

The adoption rate categories in this research were presented in the ordinal quantities from 1 to 7 and are related to adopter categories based on Rogers (2003), who classified the categories as laggards, late majority, early majority, early adopters, and innovators. The relationship between the adoption rate and the adopter categories was not a one-to-one mapping, for two reasons. First, the number of adoption rate categories in this research was seven, while the number of adopter categories in the diffusion of innovation theory was five. Second, the adopter categories in the diffusion of innovation theory described adopter categorization with a certain standard. For example, the laggards category lies at \(\bar{X} + 2sd\) interval and the innovators category lies at \(\bar{X} - 2sd\) interval in a normal distribution, while the adoption rate in this research is merely indicated by the number of semesters. Therefore, the adopter categories were simply a direction to define the adoption rate categories in a process of innovation diffusion. In fact, there was no clear-cut boundary for adoption rate categories in relation to adopter categories. Table I describes the relationship between adoption rates and adopter categories.

This study used an ordinal regression model to analyze a relationship between adopter categories and academic factors (program of study, GPA, credits earned), personal factors (age, sex, marital status, working status), and residential factor (within or outside the city). The ordinal regression required only the dependent variable as an ordinal variable. The independent variables could be an interval, ordinal, or categorical. However, to avoid too many empty cells and for convenient interpretations, the independent variables are all converted to categorical numbers, as shown in Table II.

The parameters of ordinal regression need different interpretation compared to parameters in simple linear regression. The basic concepts for interpreting parameters of ordinal regression are the concepts of odds, odds ratio, and logit (National Centre for

<table>
<thead>
<tr>
<th>First time enrolled in Bengkulu Regional Office of Universitas Terbuka</th>
<th>First time participated in the online examination</th>
<th>Number of semester before joining the online examination ((d))</th>
<th>Adoption rate ((7-d))</th>
<th>Adopters categories (as a direction for interpretations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd semester of 2013</td>
<td>2nd semester of 2013</td>
<td>0</td>
<td>7</td>
<td>Innovators</td>
</tr>
<tr>
<td>1st semester of 2014</td>
<td>1st semester of 2014</td>
<td>1</td>
<td>6</td>
<td>Early adopters</td>
</tr>
<tr>
<td>2nd semester of 2014</td>
<td>2nd semester of 2014</td>
<td>2</td>
<td>5</td>
<td>Early majority</td>
</tr>
<tr>
<td>1st semester of 2015</td>
<td>1st semester of 2015</td>
<td>3</td>
<td>4</td>
<td>Late majority</td>
</tr>
<tr>
<td>2nd semester of 2015</td>
<td>2nd semester of 2015</td>
<td>4</td>
<td>3</td>
<td>Laggards</td>
</tr>
<tr>
<td>1st semester of 2016</td>
<td>1st semester of 2016</td>
<td>5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Not participated yet until 1st semester of 2016</td>
<td>Not participated yet until 1st semester of 2016</td>
<td>6</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Table I. Relationship adoption rate and adopter categories
Research Methods, n.d). The odds express the likelihood of an event occurring relative to the likelihood of an event not occurring. For a dichotomous explanatory variable, the odds ratio is simply the difference between the odds for the base category \((x = 0)\) and the other category \((x = 1)\). Logit is the log of the odds ratio. The parameter estimate in ordinal regression as shown in Table IV is in the form of logit numbers. For instance, the parameter estimation for teachers’ category is \(-0.853\). Therefore, the odds ratio for teachers and non-teachers is \(\exp(-0.853) = 0.426\). This means that the odds of teachers achieving the higher adopter category are 0.426 times compared to non-teachers. Similarly, the odds ratio for non-teachers compared to teachers is \(1/\exp(-0.853) = 2.345\), which means the odds of non-teachers reaching the higher adopter category is 2.345 times that of teachers. In some contexts, the odds ratio could be also expressed in percentage terms, e.g. the odds ratio of adoption rate for teachers was 42 percent for non-teachers.

**Findings and discussion**

*Description of adopter category by explanatory variables*

This study dealt with 1,540 students who first enrolled in the second semester of 2013 at Bengkulu Regional Office of UT. In Table III, teachers make up 53.4 percent of the students, most notably elementary school teachers. None of the students in the teacher program was included in the innovators category. More than 50 percent of the students had a GPA of 2.9 or less in the first semester of 2016, and made up the bulk of the innovators category. Meanwhile, the students who earned 62 credits or less stood out in the laggards category with 64.0 percent.

In terms of the age group, the number of the students aged 25 or less surpassed the number of students aged over 25 in the lower adoption categories 1, 2, and 3. This led to a domination of student groups aged over 25 in the higher adoption categories 4-7, where in fact none of the students aged 25 or less were found in category 7 or classified as innovators. Classifying on gender basis, the number of female students surpassed the number of male students in almost all adoption rate categories, except for category 7. However, no such systematic difference was found for variables of marital status and working status, since the percentage between the levels in each variable converged to a percentage of the total.

As for the students’ residences, the low adoption categories 1-5 were dominated by the students residing outside Bengkulu City. In contrast, those residing in Bengkulu City
outnumbered those residing outside Bengkulu City in categories 6 and 7. This signifies that the students residing in Bengkulu City were more inclined to adopt the online examination faster than those residing outside Bengkulu City.

Model of adopter category on online examination

The parameters of the ordinal regression model for the effects of academic (program, GPA, credits earned), personal (age, gender, marital status, working status), and residential variables (within city or outside city) on the adopter categories are presented in Table IV. The model of ordinal regression proposed a significant $\chi^2$ statistic ($p < 0.0005$) to suggest that the final model led to a significant improvement upon the baseline or intercept-only model. The Nagelkerke of the pseudo $R^2$ statistics value was 16.6 percent, stipulating that the explanatory variables defined 16.6 percent of the variation among adopter categories. The model also corroborated the assumption of goodness of fit, indicated, respectively, by the $\chi^2$ test for the null hypothesis of deviance statistics that was not rejected by $\chi^2$ test at $p < 0.997$ and the $\chi^2$ test for the assumption of parallel lines that was not rejected at $p < 0.278$.

The threshold in the Table IV represents the intercept in the model of ordinal regression. All intercept coefficients were significant at $p < 0.05$ by the Wald test, which in turn would
be included in the model. The location in the table represented coefficients of the explanatory variables in the model. If the coefficient was significant at $p < 0.05$, the associated variable would be included in the model. Table IV demonstrated that all variables were significant at $p < 0.01$, denoting that there was a strong association between each of these variables and adopter categories. However, no such significant association was found for marital and working status.

As described earlier, the parameter estimate in ordinal regression as shown in Table IV is in the form of logit numbers. The logit is log of odds ratio between two categories. For a dichotomous explanatory variable, the odds ratio is simply the difference between the odds for the base category ($x = 0$) and the other category ($x = 1$). For easy interpretation of the parameters in Table IV, the last two columns of Table IV provide the exponential values of the parameters and their inverses.

### Table IV. Parameter estimation of the model

<table>
<thead>
<tr>
<th>Location ( Bs’ coefficients)</th>
<th>Estimate</th>
<th>SE</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
<th>1/Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Teachers = 0)</td>
<td>−0.853</td>
<td>0.207</td>
<td>16.928</td>
<td>1</td>
<td>0.000</td>
<td>0.43</td>
<td>2.35</td>
</tr>
<tr>
<td>(GPA = 0)</td>
<td>0.703</td>
<td>0.162</td>
<td>18.740</td>
<td>1</td>
<td>0.000</td>
<td>2.02</td>
<td>0.50</td>
</tr>
<tr>
<td>(Credits earned = 0)</td>
<td>−1.148</td>
<td>0.182</td>
<td>39.843</td>
<td>1</td>
<td>0.000</td>
<td>0.32</td>
<td>3.15</td>
</tr>
<tr>
<td>(Age = 0)</td>
<td>−0.531</td>
<td>0.190</td>
<td>7.789</td>
<td>1</td>
<td>0.005</td>
<td>0.59</td>
<td>1.70</td>
</tr>
<tr>
<td>(Sex = 0)</td>
<td>0.533</td>
<td>0.167</td>
<td>10.161</td>
<td>1</td>
<td>0.001</td>
<td>1.70</td>
<td>0.59</td>
</tr>
<tr>
<td>(Marital status = 0)</td>
<td>0.242</td>
<td>0.197</td>
<td>1.520</td>
<td>1</td>
<td>0.218</td>
<td>0.79</td>
<td></td>
</tr>
<tr>
<td>(City = 1)</td>
<td>0.915</td>
<td>0.168</td>
<td>29.695</td>
<td>1</td>
<td>0.000</td>
<td>0.40</td>
<td>2.50</td>
</tr>
</tbody>
</table>

Note: *This parameter is set to zero as it is redundant*

Online examinations at Universitas Terbuka
at UT to qualify for teaching, and evidently they are of relatively homogenous characteristics. The students of regular program, on the other hand, are of tremendously diverse kinds, from high school fresh graduates to retiring people. The differences in characteristics of such two groups brought about differences in the odds of falling into a higher adopter category to 46 percent for those in the teacher-training program.

The parameter estimate for the variable of GPA was 0.703, which was GPA > 2.9 as a baseline, suggesting that the odds of a student with a GPA of 2.9 or less to fall into a higher adopter category were 2.02 times (or 202 percent) as much as those with GPA above 2.9. The GPA data were taken in the first semester of 2016, before the inception of the online examination, which let the students decide whether to participate or not in the elapsed time – from the second semester of 2013 to the first semester of 2016. This research yet stumbled upon unexpected finding that a student with a higher GPA turned out a slower adopter. On the other hand, this finding was consistent with that of earlier research that a student with a lower GPA had a high score in expectancy and value toward the online examination (Sugilar, 2016). The study also discovered that a student with a lower GPA, in fact, needed the online examination much more as a re-examination, to improve his/her GPA. Students with lower GPA, simply put, were more likely to adopt the online examination faster. Some students, who could be categorized as laggards or late adopters, explained that they did not yet need the online examination because they were obtaining good grades for their courses. This should have supported the fact that the students who were fast adopters of the online examination were largely the students who acquired lower GPA.

It is important to note, however, that the online examination is just been another tool for measuring students’ competence in one course and assessing their failure in improving their learning so that they can learn from their mistakes. The finding that the students with lower GPA belonged to the category that adopted the online examinations faster indicates that the online examinations provided the students with the benefits of re-examination but did not lead to significantly higher scores. Therefore, there is a need for further research regarding the examination results and the students’ preparations before taking the online examinations.

The students that earned credits above 62 in the first semester of 2016 had a 3.12 times the chance to be in higher levels of adopter category as compared to those that earned 62 credits or less. In accordance with the earlier findings, this finding suggests that while the online examination did not add to the GPA significantly, the students acquired sufficient grade for credit accumulation. To conclude, although the online examination did not pull off the GPA above 2.9, it helped the students in the second semester of 2013 to gain a credit greater than 62.

**Personal factors**

Students aged 25 or less were 0.59 times slower adopters than students aged over 25. In other words, those aged above 25 were 1.7 times faster adopters than those aged 25 or less. This was contrary to the assumption that younger people, as digital natives, tended to be better adopters of communication technology. On the other hand, this finding was in line with a research result that older students demonstrate higher scores in expectancy and values toward the online examination (Sugilar, 2016). This indicated that the students considered the online examination less related to technological issues, but more associated with tools that enabled them to frequently take the re-examination. The participation in online examination was, thus, closely linked to a sense of learning seriousness. This was underlined by a research conducted by Darmayanti (1993), among others, in UT’s context, who posited that older students had seriousness and self-directed learning compared to younger students.

The odds for the female students to be part of higher adopter categories were 1.7 times that of male students; the male students had only 0.59 times the females’ possibility.
This finding extended the result of a research in online examination by Sugilar (2016) that the male students tend to participate more frequently in online examinations. The male students are consequently slower adopters than the female students, but once the male students participate in online examinations, they will participate several times more than the female students.

**Residential factor**

Residential factor refers to the place where the students resided during research period. Bengkulu Province has nine regencies and one city, which is its capital, Bengkulu City. Bengkulu Province has an estimated population of 1.828 million, of which about 310,000 are concentrated in Bengkulu City. As the capital city, Bengkulu City has better infrastructure and facilities for communication and a more heterogeneous mix of people coming from other places in Indonesia, with several foreign-born people. The people of Bengkulu City are assumed to be more cosmopolitan than those belonging to other places in Bengkulu Province.

Cosmopolitaness is a characteristic strictly associated with innovation, and people who are more cosmopolitan are earlier adopters of innovations (Rogers, 2003). Table IV reveals that students’ residential factor gives different odds for higher levels of adopter category. Those residing in Bengkulu City were 2.50 times more likely to hit the higher levels of adopter category, compared to those residing outside Bengkulu City. The city-dwelling students were thus faster adopters of the online examination.

**Conclusions and recommendations**

**Conclusions**

The results of this study showed that while program of study, grade point average, credits earned, sex, and residence had systematic effects on adoption rate categories, no such effects were found in employee status and marital status. The conclusions of this study are as follows:

- The students majoring in teacher education and training program were faster adopters of the online examination than those majoring in a non-teacher program of study. The adoption rate of the online examination among the students in the teacher program was 46 percent of the adoption rate of the students in the regular program.

- There was greater likelihood of the online examination adoption among students with lower GPA (2.9 or less) than among students with higher GPA (above 2.9), implying that students with lower GPA benefited from the online examination as a re-examination effort to improve their GPAs.

- The students with credits greater than 62 were 3.12 times more likely to achieve higher levels of adopter category compared to those with 62 credits or less.

- Students aged 25 or less had 59 percent likelihood of those aged above 25 to gain higher levels of adopter categories.

- Female students were 1.7 times more likely to pull off the higher adopter categories. The male students were thus slower adopters of the online examination as compared to the female students.

- The students that resided in Bengkulu City were 2.50 times more likely to reach higher levels of adopter categories than those residing outside Bengkulu City. These city-dwelling students, in other words, pushed on the adoption of the online examination 2.50 times faster than the non-city-dwelling students.
**Recommendations**

This study implies that the institution of ODL should accelerate the introduction of new learning resources based on diffusion of innovation modeling. Upcoming technological changes in the nature of ODL innovations are expected to extend this scope further and are expected to be integrated with students’ demographic profiles. Diffusion modeling has attempted to reflect the increasing complexity of newly introduced innovations in the adoption process. As a technological innovation, the online examination should be diffused by applying the technology acceptance model. To maximize students’ acceptance of the technological innovations, it becomes pertinent to emphasize the usefulness and easiness in terms of students’ perception. Mkhize et al. (2016) has concluded that “Technological innovation has been the pillar of success in many organisations in the knowledge age, irrespective of the organisation type” (p. 298).

In addition, it is recommended, to the developer of learning resources for the students of ODL, that all the learning resources should be developed and evaluated with students’ expectancy-value in mind (Sugilar, 2016). The following factors should be noticed: easiness, students’ ability, importance in the students’ view, possible intrinsic values of students, and cost. It is also recommended to encourage the students of ODL to have the feeling of easiness, self-confidence, and importance of using online examination, especially for the new students joining ODL. To achieve this objective, regional office staffs can guide the students to try out the online examination in the new student orientation study activities.

As this research is a case study in Bengkulu Regional Office of UT, further study is needed to compare adoption pattern of the online examination among regional offices of UT, since Indonesia consists of more than 10,000 inhabitant islands with cultural varieties. The analysis should include some variables that represent the uniqueness of each region in Indonesia. Further study is also necessary for comparing the adoption of new learning technologies among ODL students of various nations.

**References**


**Corresponding author**

Sugilar Sugilar can be contacted at: gilar@ecampus.ut.ac.id
The potential and challenges of MOOCs in Pakistan: a perspective of students and faculty

Syed Saad Ahmed, Essa Khan and Muhammad Faisal
Department of Management Sciences, Bahria University, Karachi, Pakistan, and Sara Khan
Institute of Business Administration (IBA), Karachi, Pakistan

Abstract

Purpose – The birth of massive open online courses (MOOCs) has instantly drawn the attention of scholars, academicians and learners. Millions of participants are learning through this freely accessible model of education. The purpose of this paper is to review the development of MOOCs, its characteristics and to explore its potential and challenges in Pakistan particularly.

Design/methodology/approach – The data were collected through interviews and focus group, and the respondents had completed at least one MOOC offering. This research used content and thematic analysis with the triangulation of methods and sources.

Findings – The finding of this study reflects that MOOCs are inspiring great number of learners in Pakistan despite of factors impeding the surge of e-learning. MOOCs in regional languages with better electricity and internet connectivity could be very useful for the rural areas’ people but it requires extra ordinary interest from government and academicians.

Research limitations/implications – This is an exploratory qualitative study highlighting the potential and challenges of MOOCs from the perspective of faculty and students. However, it does not incorporate the views of university officials. Similar study could consider university officials and university owner as respondents. In addition, future studies could also investigate the factors inhibiting completion of MOOCs.

Originality/value – Despite of the sharp rise of published literature on MOOCs, there is less contribution from the developing countries. This research enables us to develop better understanding of the potential and challenges of MOOCs in the social context of Pakistan.

Keywords Developing country, Human development, Open and distance education (ODE), Massive open online courses, Online learning, Free education

1. Introduction

The revolution in information technology dramatically impacted the lives of billions of human beings especially through its influential tool of internet (Nie and Erbring, 2000). This storm of wireless technology affected numbers of domains including education. Education is considered very important not only for the personal development of an individual but also for the development of any society. Schools, colleges and universities have been spearheading this endeavor for hundreds of years.

However, it is not feasible for everyone to get education from universities, especially top universities, not even for those who live in developed parts of the world because of many reasons (Tobin, 2015). Therefore, some educators conceived the brilliant idea of massive open online courses (MOOCs) to provide meaningful learning experience to anyone around
the globe for free. The term MOOC was first coined in 2008 by Dave Cormier of the University of Prince Edward Island (Miguel et al., 2013).

MOOCs are increasingly considered as an opportunity for millions of people who want access to the higher education for free or at low cost. Although academic institutions are jubilantly jumping on to the digital bandwagon of MOOCs, the majestic success of MOOCs in the recent years has also been creating panic among the university officials who believe that this free and low cost model of education would cannibalize their share (Christensen et al., 2014).

Distance learning education is not new to Pakistan. Allama Iqbal Open University was the first in Asia to start distance education in 1974 which has currently 1.3 million enrolled students (AIOU, 2016). Pakistan has got huge MOOC potential as a developing country where most of its students cannot afford to study at world’s premier academic institutions.

EdX, a joint venture of Harvard and MIT, and a premier MOOC provider have got nearly 90,000 students from Pakistan. EdX has also established international regional office in Lahore. In addition, Anand Agarwal, CEO of edX, has expressed his interest in an interview to collaborate with Pakistani universities (Ahmed, 2016a).

This research idea emanated from the personal experience and interest of researcher when I enrolled in MOOC offering of the University of Edinburgh in 2012; this was the same year when The New York Times labeled it as the year of MOOCs because of the sharp rise in the MOOC enrollments worldwide (Pappano, 2012).

Despite of the considerable boost of MOOCs and its published literature in recent years, there is dearth of research and insight from developing countries. The objective of this research paper is to review the development of MOOCs, its characteristics in general and to explore its potential and challenges in Pakistan particularly. This research study would help us to understand that what students and teachers in Pakistan think of this massive rise of this online education genie (MOOCs).

For this exploratory qualitative study, the data collection was done once from multiple groups, hence it was a cross-sectional study. Focus group was conducted with students and interviews were used to collect data from faculty members. The collected data were recorded and transcribed. We used thematic and content analysis to draw conclusions.

Finally, we concluded this paper with some useful insight that facilitate us to create a better comprehension of MOOC potential and challenges in general and particularly in Pakistan.

2. Literature review

With the economic progress of mankind and development of new technologies, new medium of imparting knowledge evolved. The emergence of MOOCs exploded in the year 2012 primarily through the US universities’ platforms such as www.udacity.com, www.coursera.org and www.edx.org (Davis et al., 2014). Many of the premier British universities joined futurelearn.com to emulate their Americans counterparts. FutureLearn was the first British MOOC platform. However, some British institutions such as the University of London and Edinburgh were already offering courses through Coursera (Parry, 2012).

These platforms are result of genuine endeavors of IT experts, scholars and professors of academic institutions (Liyanagunawardena et al., 2013). It started not long ago but its success has been phenomenal in terms of participation. Now millions of people have been learning through these platforms. In addition, Wong et al. (2016) analyzed research development in the area of distance education. They conducted content analysis of 288 research articles published in online and distance learning peer-reviewed journals. They noted that MOOCs were the most frequent keyword emerging in 18 percent of their selected articles.

Some universities are using MOOCs along with the face-to-face classes as blending teaching is increasingly evolving in this digital era (Bo Tso, 2015). It gives excellent opportunity to the students if their university is offering. However, most of the MOOCs are
based upon a freemium business model where the basic product or service is provided for free but money is charged for additional services (Dellarocas and Van Alstyne, 2013). In addition, Ming Wong (2015) explored the effect of the pedagogic orientation of MOOC platform on their courses. They examined 32 courses on the Coursera, edX, FutureLearn and openlearning. This study suggests that the platform offering MOOCs invariably influences the delivery of the MOOCs. As edX and Coursera comprehensively focused videos while FutureLearn and openlearning focused more on the participant interaction. Moreover, Kalman (2014) analyze that the distance learning education is neither new nor the pedagogic and technological approaches applied are unique. However, the concept of free education from top universities has sparked the general interest.

Altbach (2014) see the advent of the MOOCs as some form of neocolonialism where English-speaking academic cultures are dominating the MOOCs. But the millions of learners around the world are not much concerned with pedagogical philosophy and nature of knowledge they are acquiring. On the other hand, they are appreciating the MOOCs capacity to provide opportunities to everyone. Regardless of national or financial background, one can enroll in the courses through the internet access. It also opens the doors for the learners of developing countries who want to refine their skills or even start up their own ventures. According to Christensen et al. (2014), it is great opportunity for the business schools to expand into to new markets as nearly half of the MOOC offerings of Wharton School of the University of Pennsylvania have international enrollments. Furthermore, Ming Wong (2015) suggested that MOOCs help instructor to experiment his/her online teaching to large number of diverse learners. Some instructors have found MOOC experience useful for refining and enhancing their teaching skills.

MOOCs adeptly question traditional teaching method where instructor asks students whether they understand the topic and clever ones nods their heads and he/she moves on with the lecture. All five fingers are not the same, in almost every class, there are some students who hesitate to ask questions. This platform smartly addresses this crucial issue by putting video in the recorded lecture form, enabling students to access it numerous times. According to Zapalska and Brozik (2006), the instructional design in online educations must oversee different individual learning styles. Student learning is more effective if it is consistent with his/her learning styles.

Discussion forums are another exciting and informative feature of MOOCs. Students put up their questions in threaded discussions and receive numbers of responses from their fellow students dispersed around the globe (Welsh and Dragusin, 2013). While late joiners might experience problems in understanding the course requirements, and in such cases, the help from older participants was recorded very useful (Waite et al., 2013). In many courses, students from the same town and city form groups for the in-depth discussions. Li et al. (2014) also reported on the longitudinal study that the study group especially the one co-located had a pleasing experience in learning difficult topics. Moreover, self-management is sine qua non for digital learners in this information age. It will help learners to enhance learning efficiency and effectiveness in communities by developing better comprehension of valuable strategies and related concepts (Zhixian and Jianhua, 2011).

The participation certificates from top universities breed sense of achievement and encourage students to continue education. In some accredited courses, you automatically get the credits for the university education. For example, students can get undergraduate credits for some of the Coursera courses offered by Duke University, University of California, Irvine and University of Pennsylvania (Coursera, 2013). In UK, FutureLearn introduced its first MOOC having university credits, an offering by the University of Leeds that will allow student to have ten credits for its respected degree (Havergal, 2016).

Business schools need to watch closely how the genie of MOOCs is going to behave in the future. Clayton Christensen, a prominent Harvard Business School Professor envisaged that
the growth of MOOCs would shake up the foundations of business schools in the next decade and lead many of them to the bankruptcy (Kaplan and Haenlein, 2016).

The virtual platform no-pay MBA claims that it provides opportunity for those motivated individuals who ever considered MBA but fees or flexibility hampered their admission. It uses MOOCs to give a learning experience of top business schools, i.e. MIT, Harvard and Yale with the cost less than US$1,000. Their flexible program allows you to learn at your own pace. It has 14 core courses, three electives and four concentration electives. The members work together for group assignments and projects, and also meet for video discussion on a regular basis with professionals, entrepreneur and professors (Pickard, 2015).

Despite of MOOC appreciation, there has been some serious consideration, such as the assessment methodology of enrolled students has been raising questions for the MOOC credibility. For many of the courses, peer-reviewed assignments are mandatory as it is impossible for instructors to evaluate individual work. In this regard, Krause (2013) found out that some learners got commendable feedback despite of submitting substandard work. Another big challenge for the MOOCs is to ensure that the student taking the exam is the same one who is enrolled. Some experts believe that online examination with webcam could be suitable option. If their computer is locked down to use any other program or documents, this can enhance effectiveness of the invigilation (Coughlan, 2016).

Knox (2014), a Coursera Instructor interpreted the response of many participants as anxiety and burdensome because of the mass participation. Zhenghao et al. (2015) in a Harvard Business Review article claimed that over quarter of billion people had enrolled in MOOCs but expectation went down as research showed that only small proportion of enrolled students went on to complete the course (Ho et al., 2014). On contrary, Frick (2016) argues that we should not use dropout rate to claim that MOOC does not work. According to him, the other way to use MOOCs is to scan all the materials and lecture but not necessarily taking quizzes and submitting assignments. The disadvantage is evident, one would not get certificate. The advantages are that one has not got the pressure to complete the assignments and can learn at one's own pace.

Zemsky (2014) foresees that MOOCs have been losing ground as it has nothing much to offer from technological and pedagogical perspective. In addition, many business leaders think that the university graduates lack essential skills to perform the job. In this vein, MOOCs are not impacting the business world significantly because the courses are designed by the same academics who cater their own interests rather than focusing on the skills sought by industries (Weise, 2014).

In case of Pakistan, many students attempt to go to English-speaking countries particularly Australia, England and New Zealand for the higher education where securing good bands in International English Language Testing System (IELTS) is fundamental requisite for the admission. British Council which is the awarding body of IELTS certificates has launched its free course through FutureLearn, i.e. Understanding IELTS: Techniques for English Language Tests (Coughlan, 2015). This course is blessing for students who want to prepare for IELTS, as there is lack of established and standardized platform for learning English linguistic skills for higher education. Furthermore, it also prevents students to pay huge fees to substandard English language institutions.

Moreover, Agha Khan University was the first in Pakistan to launch three-week long MOOCs in 2014 on Bioinformatics of Drug Design (Abidi et al., 2016). They launched another course from the same platform in the year 2016. The organizers highlighted many factors in the context of Pakistani society for launching successful MOOC programs. First, faculty members need to leave comfort zone and review their approach in the light of global context. Second, the course design must be astutely done as course must be learner friendly with technological factor considerations.
Furthermore, the current Vice Chancellor of Information Technology University (ITU) and former Massachusetts Institute Technology Instructor Dr Umar Saif announced that ITU is working out with edX to launch online courses (Saif, 2016).

3. Methodology of research
This exploratory research was conducted for developing better understanding of the researchers (Babbie, 2008). Interpretivist paradigm with cross-sectional research design was assumed to look at the potential of MOOCs in Pakistan from people who are already aware of MOOCs. For this purpose, two set of respondents were selected: one included seven students for focus group and another comprised four faculty members for the interviews. All the individuals selected had completed at least one MOOC. Almost all of the students have done one MOOC while a student has done three MOOCs. Among faculty, all have done either eight or more MOOCs. In addition, the methodology and philosophy of qualitative study is different than quantitative. According to Marshall (1996), if the research question is adequately answered then the sample size is not an issue in qualitative study. Even for very detailed studies it could be in single figures.

The participants were requested to fill in a self-administered questionnaire requesting for personal information and brief details about MOOC already attended, before the start of the focus group or interview. The link for this questionnaire was sent to the respondents through digital means.

Following an interpretive strategy, respondents were asked to comment freely on the potential of MOOCs in Pakistan. Following were the main research questions:

- **RQ1.** How do you foresee the potential of MOOCs in Pakistan?
- **RQ2.** What will be the challenges with regard to MOOCs implementation in national and regional languages in Pakistan by Pakistani universities/professors?
- **RQ3.** What merits with regard to benefits, skills, advantages and demerits with regard to challenges, difficulties, practical and logistic problems the MOOCs will have?
- **RQ4.** What do you like/dislike about MOOCs?

These open-ended questions were asked through focus group discussion among students and through interviews from faculty. The focus group discussion last for around 1 hour 23 minutes while each interview ran from around 12-18 minutes. We had reached saturation after conducting focus group and four interviews, and according to one of the most prominent qualitative scholars, Creswell (1998), there are no rules to determine the sample size what we consider is a saturation point in the qualitative study where the new themes stop emerging. Furthermore, all discussion and interviews were recorded, transcribed and read several times to identify emerging themes or discussion falling in already identified themes. It was seen that respondents’ answers had both similarities and differences. Both thematic (Hycner, 1985) and content (Silverman, 2014) analyses were conducted as themes were already derived from the literature review and few were finalized during the analysis process. The triangulations of methods and sources were conducted – triangulation of methods, i.e. the focus group and interviews; and triangulation of sources, i.e. students and faculty.

4. Qualitative findings
Findings are presented based on both thematic (Hycner, 1985) and content (Silverman, 2014) analyses which were conducted along with the triangulation of methods, i.e. the focus group and interviews; and triangulation of sources, i.e. students and faculty.
4.1 MOOCs potential in Pakistan

Many respondents foresee MOOCs enjoying good future in Pakistan because it makes education affordable for masses. It can play positive role in the human development without huge investment in the shape of state-of-the-art buildings and it also saves considerable amount of resources. In addition, respondents also highlighted that most of the courses offered by MOOCs are from top academic institutions of the world for free. It enables us to acquire knowledge from the best professors and scholars of contemporaneous era. Apart from subject knowledge, it also boosts their general confidence level.

Couple of respondents mentioned that distance education is not new to Pakistan as Allama Iqbal Open University has been doing it for more than four decades through national television channels. As one of the respondents, Mr Khurram Adeel Sheikh said, “It is the same old candy in the new wrapping.” On the contrary, many of the respondents regarded teaching methodology of MOOCs are nontraditional and learner oriented. The lectures are recorded and uploaded in bite size videos which are easy for students to digest. They can download it and play it as many times as they want; a deed which is hard to do during real-time lectures. Moreover, it also offers opportunity to discuss topics and ideas through discussion forums. Furthermore, many of these programs are self-paced which helps to integrate every type of learners in MOOC. As one of the students in discussion forum responded, in his words, “I enrolled in MOOCs because I wanted pristine knowledge from best professors not certificates. I only downloaded the lectures and read books but never appeared in quizzes and assignments. Of course, not many people want to do quizzes in summers break.”

According to one of the university professor we interviewed, universities have to follow certain guidelines and timelines for the courses and academic programs. At times, it makes it hard for universities to incorporate all the topics in which a student could be interested to learn. Second, commencement of courses needs to have certain number of students at Pakistani universities which make students to withdraw if they do not meet quantitative requirement of the class. However, MOOCs provide learners the platform to choose the course which could be useful for the individual. Importantly, it enables learners to enjoy wide range of courses including the ones not taught in Pakistan but could see development in future.

MOOCs could be beneficial for graduates as well, as they cannot always go back to university to resume the studies. Similarly, there are some other courses as well whose teachers are not available or maybe that professor from which you are trying to get your lessons is not available. As one of the respondents uttered he wants to learn the course for artificial intelligence and if he wants to learn under the supervision of the best MIT professors so he is left with few choices. He can either join MIT but for which he might not have the IQ level or join MOOC offering where he can do it for free and also meet with hundreds and thousands of students of similar interest. This is a huge benefit which is provided by these learning platforms.

Many respondents regarded two factors financial resources and location as important factors when it comes to knowledge acquisition in the context of Pakistani society. In Pakistan, it is common to see families moving from villages and smaller towns to cities for the education of their children and most of them come from prosperous background because not everyone can afford living in the city. However, MOOCs have provided the opportunity to learn everyone at their doorsteps.

4.2 Challenges

Some respondents from the focus group concerned that MOOCs’ ability to attract large number of students becomes biggest challenge for teacher. As in many of these courses, more than hundred thousands of students are enrolled. Regardless of the competency of teacher, it is
humanly very difficult for the teacher to answer the query of hundred thousand students from around the world. So, without the real-time interaction and the questions, curiosity might be diminished particularly if concepts are not clarified. Second, students’ social interaction at the university with fellow students and teacher is very important and it refines their interpersonal and communications skills. Group assignments help them to develop team skills which are one of the most cherished set of skills at work. Moreover, one of the faculty members said in our interview that by looking at the student physically, teachers gauge their level of confidence. They closely monitor the effort of the performance. With MOOCs, they might not be able to do that because student are going to be connected via the powerpoint slides and online quizzes, consequently, evaluation of the students’ performance becomes harder.

Few respondents were skeptical regarding MOOCs success in rural area because of power failures and fluctuation of voltage. Moreover, one of respondent raised another set of concerns in the context of Pakistan that online courses unfortunately do not enjoy deserved eminence and if someone has done an online course, it is considered of lower quality without any reason. The biasness against distance learning and online learning exist in some form which we need to work on as a society. Second, some respondents were skeptical of MOOCs as they do not think employers would consider these courses because anyone can do these courses and crucially, there is no credibility of the evaluation method. However, they considered it as supplemental to one’s current knowledge and resume.

Couple of respondents mentioned that the role of Higher Education Commission (HEC – a government body responsible for facilitation and development of higher education in Pakistan) is very important in the development of MOOCs in Pakistan. If HEC promotes MOOCs at universities, it can gain recognition in a very short span of time, especially if it awards credits for the courses. On the contrary, some respondents envisaged that HEC would not do it because it might put future of the universities to questions.

4.3 The potential of local MOOC in national and regional languages
Many respondents think that MOOCs initiated in Pakistan by domestic teachers will be more useful and beneficial for the Pakistani students, as international teachers do not understand the social, religious and cultural aspects of the country.

MOOCs by Pakistani universities will have more recognition amongst their student as there would be domestic examples. As in current MOOC from western universities, most of the examples are from the institutions and organizations of developed countries which lack the social context of developing country. Doing business in USA and Pakistan are two very different cups of tea.

There are number of good-quality tuition centers which charge very high fee from O and A level students and most of them are based at cities. If some Pakistani teachers make videos of their lectures and offer it to the students, it would be a lot more convenient and beneficial for people living particularly in countryside. It may also provide opportunity to those students who cannot afford higher fee of the prestigious tuition centers.

There is great potential if MOOCs are developed in our national language Urdu and other regional languages such as Punjabi, Sindhi, Baluchi and Pashto. This will have a prospect of covering most of the underdeveloped areas of Pakistan, particularly those remote areas where women are not encouraged to go out and study. However, internet availability and connectivity is one of the greater challenges in Pakistan along with the frequent power failures. We have good internet connection only in major cities and town whereas most of the population lives in the rural areas.

5. Discussion
This research reflects that distance education programs have been firmly ingrained in our academic system for a long time. However, MOOCs are a product of information age which
uses internet to impart education and importantly, rather than operating inside national boundaries, MOOCs are global in nature, influencing people all over the world. In addition, this study notices that MOOCs got boost in Pakistan because it is free and from top academic institutions as Kalman also claimed that MOOCs did not bring revolution in the teaching methodology. Nevertheless, free courses from premier universities have enabled MOOCs to earn distinguished status.

The surge of MOOCs could see customized degree in near future where learner is free to choose from the range of courses as there are already platforms offering similar courses. The design of the MOOCs has been particularly the size of the videos and discussion forums. Similarly, Waite et al. (2013) and Welsh and Dragusin (2013) also consider it very useful as it connects all the class fellows spread around the world. Massive participation and communication makes these discussion forums like online coffee shops where you can discuss, chat and argue with your class fellows. This interaction and connection among co-located group members enables them to reap maximum learning from MOOCs.

This research shows that the enviable ability of MOOCs to attract millions of learners has also been limiting its potential as very small numbers of participants complete the courses. There could be many reasons in this regard, such as for many students, interaction with the teacher is vital for their learning style if they do not comprehend they lose interest, some students just want access to knowledge. So, they download learning material and do not participate in quizzes and assignments. Likewise, Frick (2016) believes that some students avoid evaluation because they do not like to study under pressure. They enjoy learning at their own pace.

Internet access and power failures are considered as some of the challenges for MOOCs. Pakistan is currently struggling to manage power failures as shortfall heightens up to 5,000 megawatts (Hasnain, 2017). According to the World Bank (2016) report, only 18 percent of Pakistanis are using internet, whereas Pakistan population is around 200 million with more than half of the population is under 30. Hence, there is great potential for online education as drastic changes in power and internet connectivity are imminent by dint of Chinese investment of $46 billion investment in Pakistan’s infrastructure (Ahmed, 2016b). Pakistan recorded IT export with over $2 billion last year with the growth rate of 41 percent (Baloch, 2015), this is reasonable performance considering not even quarter of the population has internet access. This research shows that students are using Coursera and Udacity courses to refine their software development knowledge and skills. Hence, MOOCs are impacting and transforming the lives of thousands of learners in Pakistan.

The support of HEC could witness the sharp rise of MOOCs culture in Pakistan but it could also endanger the future of the universities particularly business schools as predicted by prominent HBS Professor Clayton Christensen (Kaplan and Haenlein, 2016).

Hence, it could be argued that MOOCs are fine to develop basic knowledge of the concepts and theories. However, the significance of the traditional universities will remain intact in some form, seeing that some disciplines require rigorous use of expensive tools and equipment at laboratories such as disciplines of engineering and medicines which cannot be afforded by the individuals. In this vein, the augmented reality might be a game changer in the future (Castellanos, 2016a), which could help students to guide, envision nature of structures and attain deeper comprehension of the processes. This change could be expected in couple of decades. Before people used to dissect animals as part of their experiment but now there are holographic projections which can be used to explain anatomy and even pilots are being trained through it. Microsoft has also launched its holographic lenses (Castellanos, 2016b). This could be very beneficial for MOOCs and other online courses. Furthermore, idea of online laboratory is growing but this could be not deemed alternative to real one as students will not be able to develop and refine hands-on skills which are indispensable for engineers (Rasika Nandana et al., 2015).
Western MOOCs platforms are undoubtedly valuable and beneficial. However, they lack our domestic context, and believe that MOOCs by Pakistani universities could greatly help local students. On the other hand, Agha Khan University has already launched couple of MOOC programs in English language. Abidi et al. (2016) mentioned many factors universities need to consider if they are going to launch MOOCs in low-and middle-income countries. In addition, the idea of MOOCs in regional languages emphatically put forward in focus group by most of the respondents as Pakistan is much diversified on linguistic grounds and nearly 26 different languages are spoken. Nevertheless, it will require big effort and kind intentions from government as the medium of imparting education in all the premier Pakistani universities is English and most of the books and academic literature especially for higher education are available only in English language. So, this endeavor will require the translation of all the books to Urdu and regional languages. Introduction of institutional reforms to galvanize learning in the rural areas could improve the education level of the country. On the other hand, it might have some limitation as these MOOCs will be only limited to the audience who understand those regional languages.

6. Conclusion and recommendations
MOOCs intervention influenced the academic world through many conceptual and technological changes. Many academicians and scholars considered growth of MOOCs as a big threat for the traditional universities, but it has not endangered the traditional ways to imparting knowledge yet. As for many subjects, student needs close supervision, expensive equipment and state-of-the-art laboratory. However, major technological advancement in augmented and virtual reality could revolutionize MOOCs future.

MOOCs are admired in Pakistan as hundred thousands of people are already learning in the country primarily through US and British MOOCs platform. It is the best way to acquire quality knowledge at the time of heightening education cost. As discussed above, there are many sustainable benefits which MOOCs are bringing to the Pakistan society. Although limited internet access and power failures are hindering the streamlining of MOOCs, it is anticipated that in the future, structural improvement in infrastructure could positively impact the MOOC enrollments.

Launching of MOOCs in regional language is a daunting task which requires ceaseless effort from all the stakeholders. Particularly, it needs to draw special attention and interest from academicians and HEC.

MOOCs have a very low completion rate as the participants’ interest diminish with the progress of courses. Hence, for future research studies, factors responsible for inhibiting MOOCs could be investigated. The explanation of these factors along with innovative approaches could make MOOCs more influential and successful.

References


Corresponding author
Syed Saad Ahmed can be contacted at: saadahmed.buke@bahria.edu.pk

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
A multi-objective evolutionary algorithm-based soft computing model for educational data mining

A distance learning experience

Choo Jun Tan, Ting Yee Lim and Chin Wei Bong
School of Science and Technology, Wawasan Open University, Penang, Malaysia, and
Teik Kooi Liew
Department Registry, Directorate of Quality Assurance and External Relations, Wawasan Open University, Penang, Malaysia

Abstract

Purpose – The purpose of this paper is to propose a soft computing model based on multi-objective evolutionary algorithm (MOEA), namely, modified micro genetic algorithm (MmGA) coupled with a decision tree (DT)-based classifier, in classifying and optimising the students’ online interaction activities as classifier of student achievement. Subsequently, the results are transformed into useful information that may help educator in designing better learning instructions geared towards higher student achievement.

Design/methodology/approach – A soft computing model based on MOEA is proposed. It is tested on benchmark data pertaining to student activities and achievement obtained from the University of California at Irvine machine learning repository. Additional, a real-world case study in a distance learning institution, namely, Wawasan Open University in Malaysia has been conducted. The case study involves a total of 46 courses collected over 24 consecutive weeks with students across the entire regions in Malaysia and worldwide.

Findings – The proposed model obtains high classification accuracy rates at reduced number of features used. These results are transformed into useful information for the educational institution in our case study in an effort to improve student achievement. Whether benchmark or real-world case study, the proposed model successfully reduced the number features used by at least 48 per cent while achieving higher classification accuracy.

Originality/value – A soft computing model based on MOEA, namely, MmGA coupled with a DT-based classifier, in handling educational data is proposed.

Keywords Classification, Optimization, Student achievement, Educational data mining, Multi-objective evolutionary algorithm

Paper type Research paper

Introduction

For decades, educators have been troubled by questions such as “Does study habits correlate to test scores achievements?” “Why educational attainment, especially on higher educational completion rate or student retention rate is so difficult to achieve?” (Atkinson and Geiser, 2009; Nandeshwar et al., 2011), and “What is needed to improve the quality of education?” (Langstrand et al., 2015; Willman et al., 2015). We argue that by inspecting the variation in observational data on predictors (independent variables) and...
outcomes (dependent variables), our understanding of the relationship inter-played among different factors, for instance, student-teacher interactions (Allen et al., 2011) and teacher quality (Kraft, 2015) can lead to intervention instruction design that improves academic outcomes. An educational system is a complex system which comprises the pedagogy, human assets (e.g. students and teachers), supportive tools (e.g. course materials and infrastructure), social and cultural influence, and government policies, among others. At the same time, the advent of technologies, in particular the World Wide Web (as known as the internet) has also given rise to new waves in the educational domain.

In this context, distance learning spurred by the emergent of the internet, has become one of the key components in higher education. It offers adult learners the opportunity to pursue their educational dreams without the barrier of distance. Distance learning practises the delivery of course materials without needing both student and teacher to be in the same physical room. The convenience is usually made possible by computer-mediated learning or two-way interactive videos (Tan, Bong and Natarajan, 2015). As such, pedagogic strategies designed to match the new level of instructions with the level of distance learners are rapidly gaining attention among modern educators.

Additionally, the internet has also opened up new opportunities for open education in the form of open educational resources (OER). OER include books, videos, journals, articles, podcasts, lesson plans, open software, and so on at the discretion of openness and sharing (Smith, 2009). Consequently, the increase in e-learning resources and student databases leads to massive repositories of data. Although there has been considerable research on the use of data mining (DM) techniques in discovering potentially useful information from large sets of data in the fields of healthcare (Wang et al., 2013), manufacturing (Durán et al., 2010), market analysis (Fiol-Roig and Miro-Julia, 2011), and many others; only recently has researchers begun to apply them to issues of educational realm through a discipline of known specifically as the educational data mining (EDM). The plausible revelation comes from the realisation that given any context, additional hidden information may be revealed if one invests in trying. EDM involves the development, research, and application of computerised methods in detecting potential patterns from huge pools of educational data (Romero et al., 2010).

With the aggressive push towards sophisticated technology and boundary-less education, more and more factors interact within the already complex educational systems. Unfortunately, as remarked by the Science journal (Smith, 2009), we are to some extent still lacking of the understanding about the effectiveness of value-added components (e.g. OER) introduced to the educational arena. Many educational research works employ EDM techniques that rely on statistical methods (with exception a few exceptions). However, these traditional ad hoc combinations of data management tools and statistical methods are now far from adequate in analysing the vast pool of data. Therefore, it is time to seek new computing approach that can be applied to the highly non-linear, complex, and large volume of data environment. Moreover, the approach ought to be able to handle incompleteness and shortage of data. Fortunately, soft computing lends the solution. Soft computing is a collection of methodologies capable of handling imprecision and uncertainty with the aim of achieving tractable, robust, yet low-cost solutions. Such characteristics of soft computing models make them invaluably attractive when applying to DM applications.

To give an immediate example, soft computing method, in particular the genetic programming complements existing educational research by providing better insights into how student participation affects achievement (Xing et al., 2015). Furthermore, empirical research using DM technique to classify early dropout from selective Mexican high schools have been carried out with success (Márquez-Vera et al., 2016). Yet another recent application of DM technique (i.e. cluster analysis) in exploring the correlation between the students’ online interaction patterns and achievement is reported in Cerezo et al. (2016). While these approaches certainly have their strengths, the generic applications are
contested when handling multi-objective optimisation problem (MOPs). In principle, MOPs require simultaneous optimisation of conflicting objectives. As in the classification of student achievement model, one would consider improving the classification accuracy rate of the predictive results using the least numbers of available features (e.g. demographic data, social data, past examination grades, and other school-related data) from a pool of data since it is often true that the latter is more difficult to obtain or even if they do, may not be in complete set. Therefore, a robust model should hold high accuracy rate at the presence of uncertainty or data incompleteness.

This paper presents a soft computing technique, which comprises an evolutionary algorithm (EA), i.e., modified micro genetic algorithm (MmGA) (Tan, Lim and Cheah, 2013), coupled with a decision tree (DT)-based classifier, namely, C4.5 for the classification and optimisation of system. The MmGA works well in MOP context as shown by a series of previous successes (Tan, Lim and Cheah, 2013; Tan et al., 2014; Tan, Lim, Cheah and Tan, 2013; Tan, Hanoun and Lim, 2015; Tan et al., 2017). In order to evaluate the proposed soft computing model, an empirical-based case study is conducted. In particular, we aim to answer the first research question raised at the beginning of the paper. That being said, the case study attempts to uncover possibly appealing activities of students’ online interaction against test score achievement in an open distance learning institution in Malaysia. Student performance modelling is noted as the second most popular areas in EDM research (Peña-Ayala, 2014). The implications of the present work can be translated into helpful information that may assist educators in designing instructions suitable for students customisable to their learning behaviours so as to improve achievement. Kremer et al. (2013) share the similar view that technology can be used to tailor learning to student’s level of knowledge.

The remaining of the paper is organised as follows: the second section covers the background review of EDM and EA; followed by a short overview of C4.5 classifier. Third section is devoted to the description of our proposed methodology. The proposed model is first verified using a set of benchmark problems. Next, a case study for demonstrating the true efficacy of our model is described in fourth section. Finally, the paper concludes with some promising avenues for future work in fifth section.

Preliminaries

Review on EDM and learning analytics (LA)

In recent years, the potential of DM and analytics has transformed field after field. DM is the process of analysing and gleaning useful but hidden information from huge data sets (Mukhopadhyay et al., 2014a, b). DM is popular due to its ability to discover data patterns, classify objects, cluster homogeneous objects, and unveil numerous kinds of new findings (Peña-Ayala, 2014). In the educational domain, a specific form of the DM is known as EDM. EDM emerges as an approach to explore educational massive data in enhancing the educational sector. It leverages the data mined through DM to improve learning, cognition, and assessment (Sachin and Vijay, 2012).

On the other hand, LA refers to the use of learner-data in reporting and analysing models for the purpose of predicting and advising learners (Ferguson, 2012; Hwang et al., 2017). Though LA is commonly identified with EDM, they differ in terms of goals and scopes. Baker and Inventado (2014) contrast between EDM and LA in a recent review: from the technical perspective, EDM deals with the development of methodologies for analysis of learning data, while LA focusses on the interpretation of these data for optimising learning and its environment. EDM also emphasises the modelling of relationships among specific constructs, whereas LA relates the interplay among constructs from a holistic view of the system. Additionally, EDM research works concentrate mostly on the development of automated support for learners. LA, on the other hand, puts more effort in informing and empowering instructors about learners’ performance progress.
To date, EDM has taken on and extended many other related fields including text mining, machine learning, statistics, and psychometrics. Romero et al. (2013) propose a prediction model of student performance using DM techniques such as clustering and class association rules. The model has been claimed to be more representative of student groups (clusters) compared to previous rule-based model. EDM is also popularly used for student modelling (Lemmerich et al., 2011; Nandeshwar et al., 2011), which aims at characterising student by emotion, achievements, skills, learning preferences, and fulfilling individual's learning requirements through adaptation of teaching experiences. Another area of EDM application is on student assessment and evaluation, which enables student proficiency to be distinguished at a fine-grained level (Lopez et al., 2012). EDM also facilitates student feedback and support (Leong et al., 2012). More generally, EDM can be applied to educational problems with regards to emotion in context, engagement, meta-cognition, and collaboration tasks (Baker, 2014).

From the lens of LA, students' engagement and learning outcomes can be improved with proper intervention by instructors. To be effective, the intervention should be provided at the right time. LA and DM techniques are helpful in this case. A commonly used approach for discovering sequential patterns among events is known as sequential pattern mining. Chen et al. (2017) adopt the frequent sequent mining and lag sequential analysis (LSA) in order to study how learners collaborate in knowledge-building discourse. Similarly, LSA facilitates the exploration of learners' sequential patterns in other learning settings such as online interactions behaviour (Cheng et al., 2017) and problem-solving behaviour (Chiang, 2017; Hu et al., 2017). It is also not uncommon to adopt LA in analysing learners' behavioural patterns as a result of interaction with strategies or technological tools. For instance, Kizilcec et al. (2017) investigate various self-regulated learning strategies in MOOCs environment in hope of uncovering the most effective strategies and how they manifest in online behaviour. Meanwhile, Van Leeuwen et al. (2014) examine how teacher supporting tools in the context of computer-supported collaborative learning affect teacher guidance behaviour.

In general, more and more educators are turning to both EDM and LA for improving the educational outcomes. As shown by Xing et al. (2015), synthesising LA approaches and EDM techniques supplemented by genetic programming produces an effective student performance prediction model. The model has been claimed to possess higher prediction rate and interpretability compared to traditional models. Whether EDM or LA, educators can continue to benefit from the various scientific and systematic analysis methodologies available.

Review on EAs and multi-objective optimisation
Natural evolution provides a promising collection of inspirations for computational algorithms. The group of computing methodologies, which analogises the evolutionary process of biological population in finding optimal solutions to optimisation problems, is known as EA (Golberg, 1989). Generally, EA can be divided into four major classes: genetic algorithm (Holland, 1992), evolutionary programming (Fogel, 1966), evolution strategies (Schwefel, 1993), and genetic programming (Koza, 1992).

Unlike traditional methodologies, EAs are distinguished mainly by the use of a population of search space. Each member of the population corresponds to a potential solution. The quality of the solution is determined by a fitness value associated with each member. During each iteration step (generation), better fitness members receive higher chances of survival or become the parents of the next generation. Offspring which are the new population members are generated using some variation operations, like mutation and/or crossover. The evolutionary process ends after some termination criteria are met. These synergetic combinations of population-based, fitness-based, and variation-driven search have reported success in many complex optimisation problems (Tan, Lim and Cheah, 2013; Lim et al., 2015a, b, 2016; Tan et al., 2017). Meanwhile, the literature of GA runs
a long list of variance diverging from its original, yet maintaining the novelty of GA characteristics. Among the more popular ones are the micro-GA, monogamous GA (Lim et al., 2015a, 2016), island model GA, and cellular GA, to name but a few[1].

Many real-world problems are made up of performance measures (objectives) that are often conflicting in interest. They ought to be optimised simultaneously in order to achieve a trade-off. In this light, a special domain of the EA that deals with MOPs is known as the multi-objective evolutionary algorithm (MOEA). In any MOP, it is not surprising that a set of optimal solutions (as opposed to single optimum) is obtained. The optimal solution set usually consists of a number of solutions that are close in fitness according to Pareto dominance concept. As a result, comparing among the different optimal solution sets is a challenging task (Jiang et al., 2014). Various quantitative performance metrics exists in the literature of MOP for defining the optimality of different solution sets. These included the generational distance (GD) metric (Durillo and Nebro, 2011), generalised spread metric (Zhou et al., 2006), and hypervolume metric (Zitzler and Thiele, 1999).

In the meantime, the MOEA can be broadly classified into aggregation-based, indicator-based, and Pareto-based approaches. The aggregation-based approaches treat MOP as single-objective optimisation problem that can then be solved using conventional EAs after combining all its objective functions into a single weighted scalar value. However, the major shortcoming of this approach is that the scalar function and weights are critical in determining the efficiency of the algorithm. However, finding suitable weights is an optimisation problem in itself. On the other hand, the indicator-based MOEAs typically adopt selection mechanism with specific performance metric (Zitzler and Künzli, 2004; Beume et al., 2007). They have the advantage of being scalable to the number of objectives, usually four or more. However, they are generally more computationally expensive, especially when using hypervolume metric.

Finally, a representative Pareto-based MOEA approach is the MmGA (Coello and Pulido, 2005). MmGA is also an extension of the micro-GA. It has been used with great success in handling various multi-objective benchmark problems (Tan, Lim and Cheah, 2013), job-shop scheduling problems (Tan, Hanoun and Lim, 2015; Tan et al., 2017) as well as classification problems (Pourpanah et al., 2017). Even though the MmGA uses only a small size population relative to the other GA variants, it is able to achieve good convergence rate (see third section for more details). As such, this work employs the MmGA as an optimisation means. The MmGA uses GD as its performance metric.

From the DM perspective, MOEAs are popular underlying optimisation solutions for a variety of DM tasks, enumerated as clustering (Kirkland et al., 2011; Ripon and Siddique, 2009), association rule mining (Matthews et al., 2011; Martin et al., 2011), classification (Tan, Lim, Cheah and Tan, 2013; Tan et al., 2014; Pangilinan and Janssens, 2011; Pourpanah et al., 2017), and feature selection (Tan et al., 2014; Venkatachadi and Rao, 2010; Brester et al., 2014). For a complete review of the various MOEAs for DM, interested reader is referred to Mukhopadhyay et al. (2014a, b).

**Review on C4.5 classifier**

This section provides a quick overview of the C4.5 classifier, which is commonly used for generating a DT. First and foremost, a DT is a tree-like structure composed of decision rules. These rules regulate the grouping of independent variables into homogeneous zones in recursion (Cho and Kurup, 2011). DT is commonly used in acquiring information for decision making. This is in conjunction with the observation that by constructing a DT, the outcome of a set of input variables can be predicted simply by finding the set of decision rules (Pradhan, 2013). In fact, DT has been ranked as the second most popular classification methods in EDM in a recent survey conducted in Peña-Ayala (2014).
Even though there exist a plethora of DT model constructing algorithms, for instance, the chi-square automatic interaction detector DT (Michael and Gordon, 1997) and classification and regression tree (Breiman et al., 1984), this paper focuses on the use of C4.5 classifier (Quinlan, 1986) for reason of simplicity and wide applications.

C4.5 is an extended algorithm to the ID3 (Quinlan, 1986), which is based upon the Hunt’s algorithm (Hunt and Kübler, 1984). It addresses many problems that were not accounted for by its predecessor, including continuous and categorical attributes, pruning, and rule derivation. In C4.5 algorithm, a DT is built from a set of training data, \( S = s_1, s_2, \ldots \). Each sample \( s_i \) is made up of \( n \)-dimensional vector \( (x_{1,i}, x_{2,i}, \ldots, x_{p,i}) \), where \( x_{k,i} \) refers to the sample features or attribute values of class \( s_i \). When encountering continuous attributes, the algorithm simply divides the attribute values into two partitions as specified by a given threshold. In order to remove any bias of information gain, especially when an attribute has many outcome values, the C4.5 algorithm relies on gain ratio as its selection measure. Starting from the highest information gain attribute, the algorithm recurs to smaller sub-lists. In this way, the root node has the maximum gain ratio. The higher information gain attribute will be chosen for decision making (Quinlan, 1993).

Proposed model

In this work, a soft computing model to classify and optimise students’ online behaviours in a distance learning environment is presented. Students’ online behaviours as characterised by a set of web data, forms the input to our proposed model. The web data represents the frequency of students’ interactions with courses within the distance learning environment. Our aim is to classify students’ frequency of access to the learning repository against their examinations achievement at the end of a semester. Followings are elaboration of the proposed model.

Initially, a standard C4.5 classifier (Quinlan, 1993, 1996) is applied. It uses a divide-and-conquer approach to growing DTs from a set \( C \) of cases. Suppose that \( C \) fulfils a stopping criterion of decision making. The tree of \( C \) is a leaf associated with the most frequent target class in \( C \), which contains only cases of the similar target class. Meanwhile, the proportion of cases in \( X \) of \( j \)th class is identified. The uncertainty about the class for a case of \( X \) and its corresponding information gained by a test \( T \) with \( k \) outputs are computed.

Next, a specific MOEA, namely, the MmGA (Tan, Lim and Cheah, 2013) is deployed. The MmGA performs optimisation on two objective functions, i.e., maximising the classification accuracy rate (\( \alpha \)) and minimising the number of features (\( \beta \)) of classification process. Note that \( \alpha \) describes the systematic errors and measures the statistical bias in handling predictors and outcomes of C4.5 classifier processes. As articulated earlier (recall section “Review on EAs and Multi-objective Optimisation”), the MmGA is able to achieve good convergence rate as indicated by the GD metric. MmGA’s search process terminates when objective functions has reached the maximum round of evaluation or achieved convergence as measured by true Pareto. Details on C4.5 classifier as well as objective functions \( \alpha \) and \( \beta \) with relation on MmGA are presented in the Appendix.

Benchmark tests

The proposed model aims to yield a solution set, which fulfils the objective functions \( f_1 \) and \( f_2 \) such that the classification accuracy rate is maximised, while minimising the number of features during the classification stage. Prior to application on a real-world case study, we first examined the proposed model’s performance on a set of benchmark data obtained from the University of California at Irvine (UCI) machine learning repository (University of California, 2017). The benchmark data set comprises students’ achievements in mathematics and Portuguese in two Portugal secondary schools. The data attributes include student...
grades, demographic, social, and school-related features. They were collected by using both school reports and questionnaires as published in Cortez and Silva (2008).

Note only mathematical achievements, which were modelled as binary classification, but five-level classification and regression tasks were adopted in this study. We adhere to the original performance evaluation of Portugal education. That is, students are evaluated in three periods during the school year based on a 20-point grading scale (with values between 0 = lowest score to 20 = perfect score). Hence, the data set is split into three classes according to period grade, i.e., first period grade (G1), second period grade (G2), and final grade (G3). As a result, each newly created data set has originally 30 features, which correspond to variable \( x \) in the Equation (A5) for each target class G1, G2, or G3, separately. To begin, the collected grades for each class were binarised prior to classification processing: student grades were re-categorised into two groups, namely, well performed (those above or equal to score 8) and not well performed (those below score 8).

For comparison purposes, the proposed model first uses only a standard C4.5 classifier (note: in the remaining of this paper, we merely refer this model by C4.5 classifier). Subsequently, an enhanced model which incorporates the MmGA coupled with standard C4.5 classifier is deployed. It is coined as the MmGA-based classifier. The MmGA analogises the evolutionary process of biological population in finding optimal solutions for MOP. In this context, by maximising \( \alpha \) (Equation (A3)) and minimising \( \beta \) (Equation (A4)). Each member of the population corresponds to a potential solution, which is created with MmGA extended population formation. We also employ a ten-fold cross-validation method in producing the experimental results. All experiments involving both methods are repeated over 30 runs with randomised seed.

Results and discussion

Figures 1 and 2 depict the performance of the proposed model as compared to the standard C4.5 when simultaneously optimising the objective functions \( f_1(x) \) and \( f_2(x) \). Apart from a lower \( \beta \) achievement, our proposed model reported a higher \( \alpha \) relative to the standard C4.5 classifier. For completeness, the mean and standard deviation values obtained for each experiment are tabulated in Table I. Mean values marked in italics indicate best statistical significance results at 95% confidence interval under the pairwise \( t \)-test (Hall and Holmes, 2003; Götz et al., 2008) comparison.

![Figure 1](image)

**Figure 1.**
A comparison of the accuracy rate between the standard C4.5 classifier and the proposed model in the data set

Notes: (a) G1; (b) G2; (c) G3
The obvious yet encouraging results obtained inform us that our proposed model is superb in optimising the given data set using lesser number features but at the same time yielding much higher accuracy rate of classification. We attribute this to the superiority of MmGA in performing multi-objective optimisation. Consider a population of probable solutions (aka members) in our proposed model. Each member is represented as a variable $x$ following Equation (A5) and is further associated with multi-objective-based fitness values, in this case $\alpha$ and $\beta$. The quality of the member is determined by its fitness values. Like all EAs, MmGA biases members with better fitness: At each iteration step, better fitness members receive higher chances of survival or become the parents of the next generation under an elitism strategy. Offspring, or new population members, are generated using mutation, crossover, and selection operators. The evolutionary process ends with both objectives converging in MmGA nominal and outlier evolution cycles; yielding $p$ (Equation (A5)) in response to $\alpha$ and $\beta$.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Standard C4.5 classifier</th>
<th>Proposed model</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Benchmark data with G1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.6193 ± 0.0184</td>
<td>0.6472 ± 0.0170</td>
<td>9.69E−008</td>
</tr>
<tr>
<td>$\beta$</td>
<td>30.00 ± 0.00</td>
<td>15.73 ± 2.36</td>
<td>&lt; 2.2E−16</td>
</tr>
<tr>
<td>(b) Benchmark data with G2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.5499 ± 0.0189</td>
<td>0.5863 ± 0.0153</td>
<td>3.01E−011</td>
</tr>
<tr>
<td>$\beta$</td>
<td>30.00 ± 0.00</td>
<td>12.97 ± 2.46</td>
<td>&lt; 2.2E−16</td>
</tr>
<tr>
<td>(c) Benchmark data with G3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.6131 ± 0.0259</td>
<td>0.6433 ± 0.0108</td>
<td>2.12E−007</td>
</tr>
<tr>
<td>$\beta$</td>
<td>30.00 ± 0.00</td>
<td>14.03 ± 2.27</td>
<td>&lt; 2.2E−16</td>
</tr>
</tbody>
</table>

Notes: Mean ± standard deviation for 30 runs experimental results and computed $p$-values. Mean values marked in italics indicate best statistical significance results at 95% confidence interval under the pairwise $t$-test comparisons.
A case study

Satisfied with the preliminary results, let us now consider applying the proposed model to a real-world case study encompassing a Malaysian private institution of higher education with more than a decade of history in open distance education. The institution offers tertiary education to working adults via open distance learning mode. The learning environment has been catered for adult learners seeking to pursue tertiary qualifications for professional development and self-enrichment in a flexible manner. Furthermore, students and tutors come from different regions across Malaysia and worldwide.

Rather unique in its kind, the open distance learning institution provides five face-to-face tutorial classes that are spread over a period of five months to its students every semester. It also offers learning-support services via an open source learning management system (LMS). The LMS is an important platform for collaborative learning involving massive teaching-learning activities among course instructors, tutors, and students. For example, apart from the face-to-face classes, students and tutors continue to interact via video conferencing tools supported by the learning platform. Students are also free to engage in online activities such as downloading course materials, posting discussion in forums, participating in online quizzes, submitting assignments, and many more at any time anywhere convenient to them. On the other hand, instructors and tutors often play the role of system administrator in the online platform by uploading course materials, initiating discussion groups, setting up quizzes, marking assignments, answering posts, and others. It should be noted that, throughout the semester, students are generally assessed using three instruments on three periods: assignment 1 (T1), assignment 2 (T2), and final examination on the second, fourth, and fifth month, respectively.

Experimental setup

Moving on, the proposed MmGA-based soft computing model is depicted in Figure 3. Initially, data extracted from the LMS go through a pre-processing stage. It involves gathering various students’ interaction data from courses and converting their frequency into required raw data in a tabular format. Noise from the raw data are removed and transformed into a structured data format, i.e., an Extensible Markup Language file format, so that the C4.5 classifier may perform further processing. The processing stage involves employing MmGA-based soft computing model. Lastly, the output of the processing stage is made available for interpretation. In most complex systems, the interpretation may involve end-users and incorporation of other tacit knowledge to uncover the existence of any

![Figure 3. Proposed model for the EDM in an open distance learning institution in Malaysia](image-url)
possible relationship between the trends of students’ online interaction activities with the e-learning platform and their examination performance, for instance.

Students’ daily online interaction activities for every course are captured in LMS. In this study, a total of 46 courses offered in the said institution are examined. The data are collected throughout the entire semester for 24 consecutive weeks, including two weeks prior to the start of semester and two weeks after the end of semester. This contributed to 24 features, which are further grouped into two targeted classes: well-performed and not well-performed student classes corresponding to examination scores above or equal to 60 marks and below 60 marks, respectively. Note that the number of features is determined by a fixed interval of seven days. They form the input features for the classification and optimisation processes in the subsequent experimental studies carried out within the institution’s computational-based server farm.

**Results and discussion**

As depicted by Figure 4, there were initially many chosen features (i.e. weeks) resulted from the application of C4.5 classifier. After applying MmGA, a significant reduction in the number of features is observed. Worth remarking that the effect is achieved without reduction in the accuracy rates as shown in Figure 5. On average of 30 runs, there is approximately 6 per cent improvement in the accuracy rate when employing our proposed model compared to the standard C4.5 classifier. In addition, Figures 4 and 5 depict that the classification accuracy rate of the proposed model has high level of agreement with each other. This performance indicator comes from the observation of a lower number of features,
i.e., a reduction of up to 57 per cent relative to the standard C4.5 classifier. On closer inspection, the box plot distribution also reflects that the proposed model is more consistent and stable as compared to the standard C4.5 classifier since the former has narrower box and shorter tails. Reader is referred to Table II for the numerical results comprising mean, standard deviation, and p-value of the pairwise t-test comparisons. The best mean values marked in italics are statistically significant at 95% confidence interval.

To take a step further, let us examine the major determinants for the proposed model more closely. As illustrated by Figure 6, ten most prominent features (shaded in black) have been identified by our proposed model after optimisation. They represent the most significant weeks with student interaction activities that are adopted by the proposed model in classifying students achievement (recall Figure 4 and Table II). The captured interaction activities are not limited to students’ post-reply inquiry on tutorials, technical hands-on, examination-related discussions, online quizzes, manipulation of teaching-learning materials, academia-related consultation, and clarification.

At second glance, the emergent patterns unfold several interesting events. First, students are actively involved in pre-semester activities before the start of a course (weeks 1 through 2). An obvious example of such activity includes exploration of course materials by students. The trend extends towards the second week just before the commencement of a new semester, which in turn corresponds to the closing date of course registration. This has come with little surprise as students are naturally more curious and eager to know about a new course being enrolled in. But the obvious may have yet gotten the attention of educators. As evidential here, educators eager to improve students’ first perception about a course, and subsequently leading to higher motivation in continuing the course (student retention strategy) should at least invest more time in the preparation work. Early content availability and accessibility,

Table II.
Numerical results comparing the accuracy rate and number of features between the standard C4.5 classifier and proposed model on case study

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Standard C4.5 classifier</th>
<th>Proposed model</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>0.6762±0.0511</td>
<td>0.7201±0.0142</td>
<td>7.19E−05</td>
</tr>
<tr>
<td>β</td>
<td>24.00±0.00</td>
<td>10.30±2.23</td>
<td>&lt; 2.2E−16</td>
</tr>
</tbody>
</table>

Notes: Mean ± standard deviation for 30 runs experimental results and computed p-values. Mean values marked in italics indicate statistical significance results at 95% confidence interval under the pairwise t-test comparison.

Figure 6.
Adopted rate for study weeks (features) in case study
for example, would promote first positive impression and invite future interaction. A wide range of research works in cognition and social psychology attest to how initial impressions influence human interpretation of later information (Dougherty et al., 1994).

Second, weeks 7 through 10 have been recognised as other significant indicators of student performance. Inherently, plenty of practical labs preparation and T1 discussions take place throughout the second month of the semester. To educators, this is likely the best time to engage more with students in ensuring that they are well with the course. The notion arises from the assumption that increasing the two-way interactions between tutor and student will enhance both student motivation and achievement. To this end, Allen et al. (2011) find strong correlation between high-quality tutor-student interactions and improved student achievement gains.

Whereas, the third week of the fourth month and the first week of the fifth month are the remaining two indicators identified for student performance. The former reflects the presence of discussion activities following T2, while the latter arises as a result of revision activities in conjunction with final examination. The preceding makes it clear that revision is important in boosting student performance. Educators may plan revision activities on gradual basis. Revision may be treated as a form of reinforcing previously learned knowledge through practises. Intuitively, educators should pay more attention to retrieval practice in consolidating learning (Karpicke and Roediger, 2008; Karpicke and Blunt, 2011) as part of the revision process. Finally, the common activities involved in the remaining two weeks (weeks 24 through 25) may include closing of posts and reporting of web statistics. They reflect the closing of the semester.

Conclusions
Many of the developing and potential EDM research on student achievement are contended to statistical methods (with a few exceptions) and lack the knowledge in handling multi-objective optimisation problems. To this end, this paper fills the gap by proposing a soft computing model with MmGA coupled with a DT-based classifier, namely, C4.5 classification and optimisation of system. Our model has shown confident results in student achievement classification under the UCI benchmark problem and real-world distance learning case study in Malaysia by simultaneous optimising multiple objectives (performance factors), i.e., maximising accuracy rate, $\alpha$ and minimising the number of features, $\beta$. Whether benchmark or real-world case study, the proposed model successfully reduced $\beta$ by at least 48 per cent while achieving higher $\alpha$.

We believe that this work can expand access to knowledge and insight into understanding student interaction activities and their achievement based on our empirical results. It may serve as a potential platform to inform educators seeking to reform educational policy by enhancing its provision of learning-support services and create a better learning experience for the students. To this end, the results presented could be easily translated into useful information such as when and what should be done in order achieve the target research goal. In the case study, early educator preparation work, improving tutor-student interactions, and investing in retrieval practices may well improve student motivation and achievement.

This is only the beginning of a study that can lead to more elaborative outcomes for the educational arena. The results thus far could very well be true for the case study e-learning environment, but the proposed model is transferable to any optimisation and classification problems. We also plan to deploy other types of MOEAs and classifiers for other experiments in the near future. Finally, another promising work is investigation into the behaviour of the proposed model in response to the nebulous data covering different domains of interest to educators.

We need more rigorous research in the educational arena and soft computing models have opened up a new route. We believe that the future is bright and the vacuum of empirical evidence shall continue to be filled by the enthusiast research works of EDM and the alike.

MOEA-based soft computing model for EDM
Note

1. Due to space limitation, interested reader is referred to Lim (2014) for more comprehensive review of structured population GAs.

References


Appendix

C4.5 classifier
Let \( C \) denotes the number of classes and \( p(X, j) \) as the proportion of cases in \( X \) that belongs to the \( j \)th class. The uncertainty about the class for a case of \( X \), and its corresponding information gained by a test \( T \) with \( k \) outputs are derived from the following equation (Quinlan, 1993, 1996):

\[
g(X, T) = f(X) - \sum_{i=1}^{k} \frac{|X_i|}{|X|} \times f(X_i) \quad (A1)
\]

subject to:

\[
f(X) = -\sum_{j=1}^{C} p(X, j) \times \log 2(p(X, j))
\]

\[
s(X, T) = -\sum_{i=1}^{k} \frac{|X_i|}{|X|} \times \log 2\left(\frac{|X_i|}{|X|}\right) \quad (A2)
\]

Note that function \( g(X, T) \) is influenced by \( k \) and is maximal for only one case in each subset \( X_i \). Furthermore, by knowing which case the subset \( X_i \) falls, it is possible to obtained the potential information of partitioning a set of cases simply by using Equation (A2). The desirability of test is computed using the ratio of Equations (A1) and (A2).

Objective functions and MmGA
The two objective functions, i.e., \( \alpha \) and \( \beta \), of MmGA are derived as follows:

maximize \( f_1(x) = \alpha \) \quad (A3)

minimize \( f_2(x) = \beta \) \quad (A4)

subject to:

\[ F \subseteq \{ x \in \mathbb{R}^n | 0 \leq \alpha \leq 1, 1 \leq \beta \leq L \}, F \subseteq \mathbb{S} \]

\( L \) is the total of \( \beta \), and \( x = (x_1, x_2, ..., x_L)^T \) is a vector of decision variables. \( F \) and \( \mathbb{S} \) represent the feasible region and whole search space, respectively. Equation (A5) is derived to yield optimum values of \( f(x) \) based on the given \( x \). It is also known as the solution of the Pareto optimality, which denotes a Pareto optimal set with respect to the two functions \( f_1 \) and \( f_2 \):

\[ \rho = \{ f(x) \mid \forall x \in F \} \quad (A5) \]

Figure A1 further depicts the pseudo-code of MmGA, which yields the Pareto \( (\rho_{\text{true}}) \) in correspondence to true Pareto \( (\rho_{\text{MmGA}}) \). \( \rho \) has a vector of \( x^* \) that is Pareto optimal, either \( \forall x \in (f(x^*) - f(x^*)) \) or there is at least one \( i \in k \) such that \( f_i(x) > f_i(x^*) \) where \( k \) represents the number of objective functions and every \( x \in F \). As such, \( k = 2 \) in this study for optimising objectives \( \alpha \) and \( \beta \).
Corresponding author
Choo Jun Tan can be contacted at: cjtan@wou.edu.my

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com
Emerald is excited to announce a recent partnership with Peerwith, a platform that provides authors with a variety of services.

The Emerald Peerwith site can be found here: https://authorservices.emeraldpublishing.com/

Peerwith connects academics seeking support for their work with a relevant expert to get their research submission-ready. Peerwith experts can help with the following: language editing, copy editing, scientific editing, translation services, statistical support, funding application support, visuals, video, publication support, literature search, peer review and indexing services. Authors post their assignments on the Peerwith site, experts provide a quote, and the fee and conditions are then agreed upon directly between the author and the expert.

While we are not, of course, guaranteeing publication upon use of Peerwith, we hope that being able to direct academics to this resource either before submission or during the peer review process will help authors further improve the quality of their papers and increase their chances of positive reviews and acceptance.

Academics with relevant expertise can sign up as an expert on the Peerwith system here: https://www.peerwith.com/services/offer
AAOU Journal is the official journal of the Asian Association of Open Universities (AAOU).

The journal is an international, peer-reviewed, journal published twice a year. It welcomes high quality papers on topics of Open and Distance Education (ODE).

Papers submitted to the journal may be an empirical study, case study or critical literature review. It should bring readers new information, knowledge, evaluations of theories or best practices of ODE.

For more information about Emerald’s regional offices please go to http://www.emeraldgrouppublishing.com/offices

Customer helpdesk:
Tel +44 (0) 1274 783797; Fax +44 (0) 1274 783201
E-mail: support@emeraldinsight.com

Emerald is a trading name of Emerald Publishing Limited
Printed by CPI Group (UK) Ltd, Croydon, CR0 4YY

For more information about Emerald Publishing Limited please go to http://www.emeraldinsight.com

Emerald Publishing Limited
Howard Roeipe, Wagon Lane
Bingley BD16 1WA, United Kingdom
Tel +44 (0) 1274 777700; Fax +44 (0) 1274 783201
E-mail: general@emeraldinsight.com

Guidelines for authors can be found at:
www.emeraldgrouppublishing.com/aaouj.htm

Asian Association of Open Universities Journal
Indexed and abstracted by:
The British Library

Emerald Publishing Limited, Howard Roeipe, Wagon Lane, Bingley, BD16 1WA, United Kingdom
Tel +44 (0) 1274 777700; Fax +44 (0) 1274 783201
E-mail: general@emeraldinsight.com

For more information about Emerald’s regional offices please go to http://www.emeraldgrouppublishing.com/offices

Customer helpdesk:
Tel +44 (0) 1274 783797; Fax +44 (0) 1274 783201
E-mail: support@emeraldinsight.com

Emerald is a trading name of Emerald Publishing Limited
Printed by CPI Group (UK) Ltd, Croydon, CR0 4YY

For more information about Emerald Publishing Limited please go to http://www.emeraldinsight.com

Guidelines for authors can be found at:
www.emeraldgrouppublishing.com/aaouj.htm

Asian Association of Open Universities Journal
Indexed and abstracted by:
The British Library

Emerald Publishing Limited, Howard Roeipe, Wagon Lane, Bingley, BD16 1WA, United Kingdom
Tel +44 (0) 1274 777700; Fax +44 (0) 1274 783201
E-mail: general@emeraldinsight.com

For more information about Emerald’s regional offices please go to http://www.emeraldgrouppublishing.com/offices

Customer helpdesk:
Tel +44 (0) 1274 783797; Fax +44 (0) 1274 783201
E-mail: support@emeraldinsight.com

Emerald is a trading name of Emerald Publishing Limited
Printed by CPI Group (UK) Ltd, Croydon, CR0 4YY

For more information about Emerald Publishing Limited please go to http://www.emeraldinsight.com

Guidelines for authors can be found at:
www.emeraldgrouppublishing.com/aaouj.htm

Asian Association of Open Universities Journal
Indexed and abstracted by:
The British Library

Emerald Publishing Limited, Howard Roeipe, Wagon Lane, Bingley, BD16 1WA, United Kingdom
Tel +44 (0) 1274 777700; Fax +44 (0) 1274 783201
E-mail: general@emeraldinsight.com

For more information about Emerald’s regional offices please go to http://www.emeraldgrouppublishing.com/offices

Customer helpdesk:
Tel +44 (0) 1274 783797; Fax +44 (0) 1274 783201
E-mail: support@emeraldinsight.com

Emerald is a trading name of Emerald Publishing Limited
Printed by CPI Group (UK) Ltd, Croydon, CR0 4YY

For more information about Emerald Publishing Limited please go to http://www.emeraldinsight.com

Guidelines for authors can be found at:
www.emeraldgrouppublishing.com/aaouj.htm

Asian Association of Open Universities Journal
Indexed and abstracted by:
The British Library

Emerald Publishing Limited, Howard Roeipe, Wagon Lane, Bingley, BD16 1WA, United Kingdom
Tel +44 (0) 1274 777700; Fax +44 (0) 1274 783201
E-mail: general@emeraldinsight.com

For more information about Emerald’s regional offices please go to http://www.emeraldgrouppublishing.com/offices

Customer helpdesk:
Tel +44 (0) 1274 783797; Fax +44 (0) 1274 783201
E-mail: support@emeraldinsight.com

Emerald is a trading name of Emerald Publishing Limited
Printed by CPI Group (UK) Ltd, Croydon, CR0 4YY

For more information about Emerald Publishing Limited please go to http://www.emeraldinsight.com

Guidelines for authors can be found at:
www.emeraldgrouppublishing.com/aaouj.htm

Asian Association of Open Universities Journal
Indexed and abstracted by:
The British Library

Emerald Publishing Limited, Howard Roeipe, Wagon Lane, Bingley, BD16 1WA, United Kingdom
Tel +44 (0) 1274 777700; Fax +44 (0) 1274 783201
E-mail: general@emeraldinsight.com

For more information about Emerald’s regional offices please go to http://www.emeraldgrouppublishing.com/offices

Customer helpdesk:
Tel +44 (0) 1274 783797; Fax +44 (0) 1274 783201
E-mail: support@emeraldinsight.com

Emerald is a trading name of Emerald Publishing Limited
Printed by CPI Group (UK) Ltd, Croydon, CR0 4YY

For more information about Emerald Publishing Limited please go to http://www.emeraldinsight.com

Guidelines for authors can be found at:
www.emeraldgrouppublishing.com/aaouj.htm

Asian Association of Open Universities Journal
Indexed and abstracted by:
The British Library

Emerald Publishing Limited, Howard Roeipe, Wagon Lane, Bingley, BD16 1WA, United Kingdom
Tel +44 (0) 1274 777700; Fax +44 (0) 1274 783201
E-mail: general@emeraldinsight.com

For more information about Emerald’s regional offices please go to http://www.emeraldgrouppublishing.com/offices

Customer helpdesk:
Tel +44 (0) 1274 783797; Fax +44 (0) 1274 783201
E-mail: support@emeraldinsight.com

Emerald is a trading name of Emerald Publishing Limited
Printed by CPI Group (UK) Ltd, Croydon, CR0 4YY

For more information about Emerald Publishing Limited please go to http://www.emeraldinsight.com

Guidelines for authors can be found at:
www.emeraldgrouppublishing.com/aaouj.htm

Asian Association of Open Universities Journal
Indexed and abstracted by:
The British Library

Emerald Publishing Limited, Howard Roeipe, Wagon Lane, Bingley, BD16 1WA, United Kingdom
Tel +44 (0) 1274 777700; Fax +44 (0) 1274 783201
E-mail: general@emeraldinsight.com

For more information about Emerald’s regional offices please go to http://www.emeraldgrouppublishing.com/offices

Customer helpdesk:
Tel +44 (0) 1274 783797; Fax +44 (0) 1274 783201
E-mail: support@emeraldinsight.com

Emerald is a trading name of Emerald Publishing Limited
Printed by CPI Group (UK) Ltd, Croydon, CR0 4YY

For more information about Emerald Publishing Limited please go to http://www.emeraldinsight.com

Guidelines for authors can be found at:
www.emeraldgrouppublishing.com/aaouj.htm

Asian Association of Open Universities Journal
Indexed and abstracted by:
The British Library

Emerald Publishing Limited, Howard Roeipe, Wagon Lane, Bingley, BD16 1WA, United Kingdom
Tel +44 (0) 1274 777700; Fax +44 (0) 1274 783201
E-mail: general@emeraldinsight.com

For more information about Emerald’s regional offices please go to http://www.emeraldgrouppublishing.com/offices

Customer helpdesk:
Tel +44 (0) 1274 783797; Fax +44 (0) 1274 783201
E-mail: support@emeraldinsight.com

Emerald is a trading name of Emerald Publishing Limited
Printed by CPI Group (UK) Ltd, Croydon, CR0 4YY

For more information about Emerald Publishing Limited please go to http://www.emeraldinsight.com

Guidelines for authors can be found at:
www.emeraldgrouppublishing.com/aaouj.htm

Asian Association of Open Universities Journal
Indexed and abstracted by:
The British Library
1 A design-based approach to support and nurture open educational practices
Shironica P. Karunanayaka and Som Naidu

21 Learning analytics in higher education: an analysis of case studies
Billy Tai Ming Wong

41 Student characteristics and knowledge on ODL concepts at first registration: a case study from OUSL
K.A.J.M. Kuruppuarachchi and K.O.L.C. Karunanayake

52 Analyzing students online learning behavior in blended courses using Moodle
Rosalina Rebucas Estoilo and Rodolfo C. Catanta Raga Jr

69 Quality assurance in learning material development at OUM
Nazrai Ahmad Zabidi, Tai Kwan Woo, P. Rajesh Kumar, Mansor Fadzil and Syarifah Hidayati Syed Husain

82 The online examinations at Universitas Terbuka: an innovation diffusion viewpoint
Sugilar Sugilar

94 The potential and challenges of MOOCs in Pakistan: a perspective of students and faculty
Syed Saad Ahmed, Elisa Khan, Muhammad Faisal and Sara Khan

106 A multi-objective evolutionary algorithm-based soft computing model for educational
data mining: a distance learning experience
Choo Jun Tan, Ting Yee Lim, Chin Wei Bong and Teik Kooi Liew