Learner Behaviour in HSC Study Lab

Eamon Vale, MLS&T

Macquarie University

Author Note

Eamon Vale,

The School of Education

Faculty of Arts

January 23rd, 2020

Acknowledgements

I would like to take the opportunity to thank my wife (Emi) for her love and support, without which none of this would have been possible. I would like to thank my friend and mentor Ben for his advice and encouragement.

However, most of all I would like to acknowledge the guidance, understanding and above all patience of my supervisor Garry Falloon who sometimes led me, and other times dragged me through the process of conducting and writing this Master of Research Thesis. Thank you for all your help, advice and feedback.

Contents

Statement of Originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

______7/11/2019______________________

Abstract

Online learning with no direct teacher instruction, facilitation or student support is increasing within education. These autonomous learning environments present challenges for educators including how to identify student learning behaviours as well as how to offer support to isolated and/or disengaged learners.

This research project investigated the efficacy of learning analytics as a tool for identifying and evaluating student behaviours associated with active learning within an online learning environment. The study harvested data from a digital video player embedded in a popular K12 online science program and applied a learning analytics method in order to analyse active viewer behaviour. This data was supplemented by additional information from a questionnaire in order to provide greater detail as well as validate the findings of the learning analytics method. The study concluded that the learning analytics' method was able to identify active learning and that learning analytics has a role to play in providing information on learner activity in autonomous learning environments. However, as shown in the results of the questionnaire, the levels of active learning the LA method could identify were not a complete picture of learner behaviour and viewing strategies.

It is anticipated that the results from this study will provide guidance on the implementation and use of learning analytics in video-based learning for identifying active learning.

Keywords: Learning analytics, video-based learning, visualization tools, interactive learning environments, active learning.

Introduction

This study evaluated the use of Learning Analytics (LA) to identify behaviours associated with active learning in video-based lessons in HSC (High School Certificate) Study Lab. It is claimed that tools such as LA can inform administrators and potentially trigger interventions in independent online learning environments in which traditional facilitator support or oversight is limited (Pardo, 2014). This study is an exploratory research project that deliberately employed tools ubiquitous in video-based learning and therefore available to most course administrators. It initially focussed on evaluating the effectiveness of the LA method and establishing a link between the data and the underlying behaviours, with the intention of leading to more substantive research. Future research will extend the study to new environments and richer data sets in order to add to the validity and reliability of the current results. The goal of future research in the area of LA in online learning will be to establish a tool that efficiently informs on student learning behaviours within online learning environments and can provide this information to course designers and facilitators in a timely and effective manner.

In their article on the 'Impact of Technology and Theory of Instructional Design since 2000' Warren et al., (2014) identify the most recent paradigm shift in education as being the 'Age of Learning Environments' as defined by mobile learning, multiuser virtual environments, and games and simulations. These environments meet the needs of a changing demographic of student, with an increase in non-English (and English as a second language) speakers, and with the requirements for a new type of economy; that of the post-industrial information age (Warren et al., 2014). In response to these pressures online learning has grown exponentially over the last 10 years, and is an integral part of the delivery model of most educational institutions, with 90% of institutions within the US now offering some form of online education (Bowers & Kumar, 2015). According to the Sloan Consortium cited in Bowers et al., (2015, p. 28) the growth in enrolments for online courses has exceeded the growth in enrolments for traditional brick and mortar courses. This finding is supported by Allen and Seaman (2015), who also identified online enrolments as greatly outpacing those of the overall growth in higher education.

It is accepted that under the right circumstances, technology benefits learners (Winn, 2002) and the move to online learning has been identified as providing a number of advantages over more traditional modes of teaching (Spector, 2001). Online learning offers advantages such as convenience, flexibility, and access over traditional 'face to face' models (Bowers & Kumar, 2015). Students now often view online learning as their preferred model and as such universities and other education programs are compelled to compete for online students (Fekula 2010, as cited in Fedynich, Bradley, & Bradley, 2015). While it is apparent that there is a continued increase in the move towards online learning, it has not followed that online learning is a better model in terms of student outcomes, and as such, it has not removed the role of learning design but rather increased its necessity (Spector, 2001). Sims, Dobbs and Hand (2002) note that effectiveness of online learning and outcomes for students have been mixed, and a number of authors have pointed to the high attrition rates in online courses as a cause for concern (Bowers & Kumar, 2015; Kizilcec & Halawa, 2015). With the continued demand for increased online learning options and a mixed report in terms of student learning outcomes, it is not surprising that a consistent message from universities has been an eagerness to understand student learning behaviours in the online setting (Coffrin, Corrin, de Barba, & Kennedy, 2014). However, although research in online learning is increasing, widespread belief in the efficacy of online learning and educational technology is still generally intuitive, and the support for it is primarily driven by a desire to be competitive in the global education economy rather than proven educational outcomes (Harris & Walling, 2014).

Online learning environments require regular evaluation and assessment to ensure effective quality control (Sims et al., 2002). For a long time, this has meant evaluating online learning programs using time consuming and invasive data collection methods. However, digital tools are beginning to present new possibilities for the collection and observation of interactions occurring within these environments (Pardo, 2014). These digital tools, along with an increase in data quantity, advances in computing, and tools for analysis (Baker & Inventado, 2014), have helped to create a whole new field of research within the learning sciences. This new field of research has been termed 'learning analytics' (LA) - sometimes 'learning and knowledge analytics' (LAK), and its aim is to use learner data to develop a greater understanding of learner behaviour (Verbert, Manouselis, Drachsler, & Duval, 2012). A commonly cited definition of LA comes from the 1st International Conference of Learning Analytics: "(T)he measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs"

(Siemens & Long, 2011, p. 34). LA may present an opportunity to identify student engagement and success, as well as overall quality within online learning in an efficient and cost-effective manner. This area of research has developed in response to both the opportunities and challenges afforded by the vast increase in educational data produced by these new learning environments (Behrens & DiCerbo, 2014).

This study evaluated the efficacy of LA at identifying student learning behaviours within the HSC Study Lab online learning program; specifically, how students engaged with the video-based learning artefacts. In order to identify and investigate student engagement, however, it is necessary to adopt a framework to define that engagement (Ma, Han, Yang, & Cheng, 2015). Therefore, this study evaluated forms of engagement associated with active learning and adopted the ICAP framework as the model for defining active learning. The ICAP framework was developed by Chi and Wylie (2014) and provides a model of active learning defined by overt engagement behaviours first described by Bonwell and Eison (1991) and Chi (2009), before being further differentiated into the ICAP framework (Chi & Wylie, 2014). ICAP is an acronym made up of the (sub)modes of active learning and are listed as Interactive engagement, Constructive engagement, Active engagement and Passive engagement (Chi & Wylie, 2014). The ICAP framework makes a number of assumptions supported by experimental studies and a meta-analysis of existing studies, that the behaviours are a reflection of the learner's underlying cognitive engagement (Chi et al., 2018). By adopting the framework's definition of active learning, and evidence in support of the underlying cognitive engagement, this study could infer that the activity, captured in learner data, could be said to be 'active' (or otherwise) if it matched that defined by the framework. Finally, within LA there is an emerging focus on K-12 students as well as an argument that learning environments incorporating several supporting digital technologies are valuable areas of research (Warren et al., 2014). Mayer (2017) also identified K12 classrooms as the latest location for the adoption of eLearning where, he claimed, online learning environments are now common. As such, it is both timely and relevant that this study investigated an online learning environment for K-12 students.

Research Questions

The study applied an LA method to investigate evidence of active learning in patterns of student behavior within the HSC Study Lab. A secondary method comprising a questionnaire, was employed to provide additional insight into student behaviour and motivations, as well as the effectiveness of learning analytics to identify active learning in video-based learning.

Therefore, the research questions the study explored were as follows:

1. To what extent do students participate in active learning behaviours in HSC Study Lab?
2. To what extent is learning analytics an effective tool for identifying patterns of behaviour associated with active learning in HSC Study Lab?

Research associated with the areas of online learning environments, video-based learning, active learning and learning analytics are of relevance.

The Online Learning Environment

The online learning environment in which data for the study was harvested was HSC Study Lab. HSC Study Lab is an online learning program developed by Macquarie University for the purposes of helping improve learning outcomes for students in years 11 and 12 of high school within the subjects of Physics, Chemistry and Biology. It exists in a digital ecosystem through which learner behaviour (in the form of trace data) can be observed, recorded, and analysed. Designed around an 'anywhere, anytime' learning vision, learners are completely independent within the environment. However, as such it is difficult for course designers to evaluate student learning and interaction with the program content. There are a number of contributing factors to the success of online learning environments including the design of learning objects within the program (specifically if they are designed according to the cognitive theory of multimedia learning (Mayer, 2017) as well as flexibility, engagement, and quality (Cashion & Palmieri, 2002). However, in order to inform on the efficacy of the program tools and design, it is necessary that educators understand how students are engaging and performing within the learning environment (Ali, Hatala, Gasevic, & Jovanovic, 2012).

Within HSC Study Lab all content is delivered using video-based tutorials, experiments and animations; the program also features many lab type simulators and traditional recall style assessment activities (quizzes). Little is known about how learners interact with the objects in these online learning environments (Winne, 2018). As the video-based tutorials and animations are the main vehicle for content delivery any attempt to understand how the learners are interacting with the content of the program requires an understanding of how they're interacting with the videos. However, there is currently no built-in system for assessing how students are interacting with the videos, which makes it difficult to evaluate such for quality and/or identify if there are struggling or disengaged learners. Research reveals that learners benefit from guidance within these environments, and that for learners with low levels of prior knowledge, a lack of guidance can be detrimental to the learning process (Mayer, 2004; Scheiter, 2014). Furthermore, a number of studies have revealed that learners who are struggling either do not use or misuse help functions so are particularly vulnerable in instructorless environments (Aleven, Stahl, Schworm, Fischer, & Wallace, 2003).

Literature Review

An Introduction to Active Learning

Often described as a process by which a learner is an active participant in their own learning (Chi, 2009) active learning is a key learning strategy for student success. A review of the literature reveals a number of similar yet differing conceptual models for active learning. Zimmerman (1990) described active learning as learning in which learners are "metacognitively, motivationally and behaviourally active participants in their own learning" (Zimmerman, 1990, p. 4). Bonwell and Eison (1991) further elaborated with a model of active learning defined by overt behaviours that were both observable and measurable. Within their model active learning was separated into distinct modes of activity such as, reading, writing or discussing (Bonwell & Eison, 1991). Bonwell and Eison (1991) state that these behaviours must be meaningful i.e., they must be linked with 'higher-order' cognitive behaviours like 'analysis' and 'synthesis' in order to be considered active. According to Bonwell and Eison (1991) behaviours associated with active learning hold several advantages over passive learning behaviours and should be encouraged and/or modelled in the classroom. Furthermore, Bonwell and Eison (1991) state the use of these behaviours is vital as they have a powerful impact on students' learning and that studies reveal students prefer active learning over traditional lectures.

Active learning is often contrasted with passive learning, which was based on behaviouristic principles that saw learning as an input-output model in which students learned automatically based on environmental stimuli (Bandura, 2001). It is now known that this is not the optimal way to learn and according to Scardamalia and Bereiter (2006), much of the effort within education during the 20th century was a move from this transmission model of learning to what is now generally called active learning. van Hout-Wolters, Simons and Volet (2000) state that active learning is learning in which the learner makes decisions about the process of how they learn as well as the extent to which they're challenged to use their mental abilities. Active learning is a more attractive form of learning than passive learning and learners are more motivated and interested when they are in control of their own learning (van Hout-Wolters et al., 2000). Taken collectively these models provide a general understanding of active learning in which learners are physically, motivationally and cognitively engaged in the process of their own learning.

Active learning and constructivism

Active learning has a long tradition as part of the constructivist paradigm of teaching and learning (Mayer, 2004). In fact it could be argued that Vygotsky (1980) spoke specifically of what would later be termed active learning, when he described the co-construction of higher psychological function through the use of tools and signs with psychological activity (Vygotsky, 1980). According to Mayer (2004) although constructivism takes many forms its "underlying premise is that learning is an active process" (Mayer, 2004, p. 14). Active learning according to Mayer is often defined by behaviourally active activities like group discussions, hands-on activities, and interactive games, whereas passive learning is defined by reading books, listening to lectures or viewing on-line presentations. However, Mayer (2004) describes how a constructivist view of learning in which a learner is cognitively active has been misinterpreted or mis-applied as a theory of teaching in which the learner should be behaviourally active. This he terms the 'constructivist teaching fallacy' (Mayer, 2004). According to Mayer, the most important determinant of active learning is cognitive activity (Mayer, 2004). This view of the dualistic nature of active learning is supported by other constructivist theorists like Bandura. Bandura describes the need to address "the formidable explanatory challenge for a physicalistic theory of human agency and non-dualistic cognitivism" (Bandura, 2001, p. 4) and describes a model for active learning that is motoric (physical)

- action matched with underlying cognitive processes. Bandura (2001) stated that there was significant support for this model within research on brain development.

Self-regulated learners have been described as active learners who manage their own learning (Winne & Perry 2000, as cited in Azevedo, Guthrie, and Seibert 2004) . Like active learning, self-regulated learning is a constructive process, in which learners monitor, regulate and control their behaviours in meeting their learning goals (Azevedo et al., 2004). van Hout-Wolters, Simons, and Volet (2000) define active learning as the way in which a learner decides about aspects of the learning process itself, while also being the extent to which a learner is “challenged to use his or her mental abilities while learning.” (van Hout-Wolters et al., 2000, p. 21). They describe a dual nature of active learning that has to do with either decisions about learning or making active use of thinking. This first form of active learning (decisions about learning) is termed ‘self-directed learning’ by van Hout-Wolters et., al (2000) and is very similar to a definition of self-regulated learning (SRL). SRL is a crucial skill within online learning, specifically in autonomous learning environments like HSC Study Lab. It has been argued that with the explosion in the availability of information – specifically due to the rise of the internet, it is essential that students learn the skills of self-regulation and that teachers need to design instruction to teach the strategies of filtering, selecting, and processing information (Boekaerts, 2017). Typically, as students move from secondary to tertiary education they move from a more prescriptive and supportive learning environment into one that requires greater independence and time management, and SRL becomes a critical skill in this new environment (Colthorpe et al., 2019).

Active Learning in Online Learning Environments

Educational computing was popularly thought to encourage active learning, however, research has revealed that it instead tends to support knowledge reproduction – a phenomena associated with passive learning (Scardamalia & Bereiter, 1993). Wiseman, Kennedy and Lodge (2016) state that supporting and facilitating student engagement in online learning is the main challenge for educational technologists. Wiseman et al., (2016) use the term ‘student engagement’ instead of active learning but considering they define engagement as being a student’s interaction with a learning task and determined by the three dimensions of ‘cognition’, ‘affect’ (motivation) and ‘behaviour’, it fits the same broad model as the earlier definitions and thus can be considered (at least generally) synonymous. According to Wiseman et al., (2016), autonomous student learning is becoming increasingly important as digital (online) learning environments have reduced academic support, and as a result, the necessity for student self-motivation and independence has increased. Within online learning environments it is difficult for course designers and/or administrators to determine how learners are interacting with the online tools, and whether or not meaningful learning is occurring (Ruipérez-Valiente, Muñoz-Merino, Leony, & Kloos, 2015; Stepanek & Dorn, 2017). Furthermore, although the body of literature on student engagement (active learning) is broad, how it is operationalised in different contexts such as online learning is less clear (Wiseman et al., 2016). Finally, active learning behaviours in video-based learning have been found to have beneficial learning outcomes with an increase in student performance compared to passive viewing behaviours (Kleftodimos & Evangelidis, 2016). Studies by Stepanek and Dorn (2017) and Ma et al., (2015) also found that ‘active learning’ specifically related to video-based learning, was a good predicator of success. As such, this study investigated the use of LA for identifying student active learning behaviours in video-based learning within an online learning environment.

Active Learning and the ICAP Framework

With the long tradition of active learning in the literature and the variation in its definition, it became necessary to adopt one working definition as a determining framework. For the reasons laid out in this section, this study chose the ICAP framework for active learning. The ICAP framework divided and ranked active learning by (sub)modes of engagement, which were mapped to overt motoric behaviours. The ordering of the separate modes of engagement are: Interactive engagement, Constructive engagement, Active engagement and Passive engagement (being the absence of active learning) with the terms listed coming together to form the acronym ICAP and expressing a hierarchy of I>C>A>P. Chi and Wylie (2014) argue that this hierarchy of engagement correlates to a matched level of learning, with Passive resulting in the lowest and Interactive the highest learning outcomes; this is described as the ICAP hypothesis (Chi & Wylie, 2014). As the ICAP framework differentiates active learning into a hierarchy of engagement levels it is possible to categorise the behaviour and make assumptions as to the underlying cognitive engagement.

Active Learning and Engagement - Nomenclature

The ICAP framework separates active learning into modes of engagement; 'Interactive engagement', 'Constructive engagement', 'Active engagement', and 'Passive engagement' and as such the behaviours may be referred to as active learning or Active, Constructive engagement etc. depending on the specific behaviour being discussed. In order to reduce the potential for confusion, when specifying a mode within the framework of active learning the mode has been written in its capitalised form e.g. Interactive engagement, Constructive engagement, or Active engagement'. Please note that all these modes fall under the umbrella term and are evidence of active learning.

Knowledge Change Processes

The ICAP framework assumes that students' overt behaviours and associated outputs can determine their level of cognitive engagement (Chi et al., 2018) and that this engagement produces distinguishable 'knowledge-change processes'(Chi et al., 2018). The differentiated modes of activity can be mapped against levels of cognitive engagement (and thus learning); with Passive corresponding to 'minimal understanding', Active to 'shallow understanding', Constructive to 'deep understanding', and Interactive as 'deepest understanding' (Chi & Wylie, 2014). In turn each level of cognitive engagement leads to specific 'knowledge-change' processes (Chi & Wylie, 2014). So Interactive > Constructive > Active > Passive lead to the knowledge change processes of Co-Infer > Infer > Integrate > Store (Chi & Wylie, 2014). These processes and their definitions are provided in Table 1.

Table 1

ICAP Framework listing the knowledge change process and its definition

Knowledge change process	Definition
Store (Passive)	is defined as new information being stored but in an isolated fashion i.e., not integrated with prior knowledge.
Integrate (Active)	is defined as new information that activates relevant prior knowledge and is integrated with activated prior knowledge.
Infer (Constructive)	is defined by new information being integrated with activated prior knowledge, with the additional stipulation that the new information is inferred from this integration with activated prior knowledge i.e. it is constructive.
Co-infer (Interactive)	is defined by the learner inferring new knowledge from activated and integrated knowledge, but with the additional stipulation that it is an iterative and co-constructive process i.e. it is the result of a collaborative exercise (Chi & Wylie, 2014).

The ICAP Framework Applied to Video Based Learning

When learners engage in behaviours like reviewing sections of a video, such behaviours can be interpreted as emphasising certain parts of a learning material (Chi et al., 2018). An assumption regarding underlying knowledge building can also be made, in this case that the learner is building a partial schema or mental model of the information presented (Chi & Wylie, 2014). This schema provides the learner with a scaffold which they can return to and develop through methods like gap filling (Chi & Wylie, 2014). In video-based learning, Chi and Wylie (2014) state that engagement can be said to be ‘active’ if some form of overt motoric or physical manipulation takes place and specifically identified: ‘pausing’, ‘playing’, ‘fast-forward’, and ‘rewind’ as video manipulation. Dodson et al., (2018) adopted the ICAP framework in their study, and further defined it for video-based learning to include: browsing, searching, pausing, changing play back speed and re-watching video content, while passive viewing was defined as watching a video linearly without interaction (Dodson et al., 2018). This study applied the extended behaviours as defined by Dodson et al. (2018) in its framework.

Potential for Misalignment Between Observed Behaviours and Underlying Motivations

For the outcomes from this study to be meaningful it is important to establish that the learner motivations and/or intentions behind any identified behaviours align with the identified behaviours. This is because, as discussed in the literature, active learning is not simply a question of behaviour, but also underlying intention. The determining intention for active learning according to Chi and Wylie (Chi & Wylie, 2014) is ‘focussed attention’ i.e., that the activity causes the student to focus attention on the activity. As an example of action matching intention they provide “pausing and rewinding parts of a video tape” (in order to review certain selected parts of the tape)” (Chi & Wylie, 2014, p. 222). This would contrast with the pausing and rewinding due to a fault in the video or video player and/or otherwise unrelated to the information being viewed or sought by the learner.

Learning Analytics

Universities and other educational institutions are increasingly investing in online learning and a third of all enrolments in the US are now in online courses (de Freitas, Morgan, & Gibson, 2015). This move from classrooms to online learning environments has created what DiCerbo and Behrens (2014) have termed a ‘digital ocean’ of trace data which in turn has created a new and emerging area of research in learning technology known as ‘Learning Analytics’ (Verbert et al., 2012). With the increase in student generated data stored in online learning programs the analysis of this data (LA) is increasing as a field of research (Stepanek & Dorn, 2017). Learning Analytics (LA) is a new field in the learning sciences developed to exploit the potential of student data to provide insight into learner behaviours and has been identified as “the most critical factor influencing higher education” (Ma et al., 2015, p. 27). The use of LA in education has grown in recent years for four primary reasons: a substantial increase in data quantity, improved data formats, advances in computing, and increased sophistication of tools available for analytics (Baker & Inventado, 2014). Using the methodologies associated with LA, patterns can be revealed in learner behaviour (Baker & Inventado, 2014), which in turn have the potential to reveal motivations behind learner actions. Although there has been much recent research on its effectiveness as a tool for instructors and teachers, there are areas where further work is needed. According to Roll & Winne (2015) further research is required in LA as it relates to self-regulated learning, in capturing context, in working with large data sets, in the challenges presented by social learning and in being able to convey analytics to learners in a meaningful way.

Baker and Inventado (2014) argue that collecting data in real-time is often expensive and disruptive, while a data based prediction model is preferable as it is non-intrusive. Pardo (2014), stated that due to the time consuming and invasive methods previously employed, collecting feedback and using it to improve learning has had a “fairly reduced impact” (Pardo, 2014, p. 15). Pardo notes that traditional educational research is often highly invasive while LA offers the possibility of collecting observations with little to no user interaction (Pardo, 2014). Winne (2017) further contrasted traditional random controlled trials (RCTs) with the advantages seen in using big data and LA. These advantages include data being gathered during the process of learning and using the everyday tools of the program, so there is less of a disconnect between the experiment and the real-world environment (Winne, 2017). The study size can also increase from a random sample to equal the population of learners, so it is claimed that it is easier to draw inferences from the sample to the actual population (Winne, 2017). Finally, due to the ‘precision and fullness’ of data, it is argued that replication becomes a comparatively simple process (Winne, 2017).

Pardo (2014) states that an advantage of LA is that the information it provides on student behaviour can form the basis of real-time interventions. Furthermore, that it has potential to provide information on learners in an online learning environment where such an operation is otherwise not available, given the size of student cohorts (Pardo, 2014). An example of the type of intervention given by Pardo (2014) included detecting students who are at risk of abandoning a course and the use of this information to alert teachers, or otherwise take actions, to support learners. HSC Study Lab has just under 20,000 users and traditional research methods such as questionnaires and surveys would only be able to reach samples of that population. Such research methods are also invasive and time intensive so it was worth exploring whether LA could provide a viable alternative method of evaluation. The potential to course administrators of an LA method is that such information could inform course and/or video design as well as leading to the development of learning interventions. Finally, Baker and Inventado (2014) state that a positive feedback loop has been observed in LA with discoveries leading to changes in practice so there is evidence in support of the contention that this study could lead to improvements in course design or facilitation.

A LA method for identifying learner behaviour in video-based learning is congruent to the environment, non-invasive, and deals in large amounts of highly contextual data, and research reveals that inferences can be made from the meaning of that data (Pardo, 2014). However, for all the potential and claims made of LA there is not yet consensus on the approaches or the ultimate benefits, with Zilvinskis et al., (2017) stating that even after successful implementation of an LA study, there appears to be little evidence of substantial improvements in learning. Furthermore, a meta-analysis conducted by Viberg et al., (2018), found that only 9% of the LA studies they reviewed, reported improved learner outcomes. With that in mind there is plenty of opportunity for this study to contribute to and inform on the potential of LA as an effective method within online teaching and learning.

LA in Video-Based Learning

A review of the literature has revealed that LA is a research area in its infancy (Pardo, 2014) and this appears to also be true of video-based learning, with current literature lacking sufficient understanding of how students engage with video content (Dodson et al., 2018). This provides both a challenge and opportunity for this study as both its primary method of analysis (LA) and research environment are comparatively under-researched. Chorianopoulos, Giannakos, Chrisochoides, & Reed (2014) commented that there has been little and scattered research into the use of learning analytics to understand learner experience in video-based learning. They claimed that there is limited understanding of the effectiveness of video, and of how students learn from video lectures, concluding that the research area is ‘embryonic’ (Chorianopoulos et al., 2014). According to Lodge et al., (2017) the lack of empirical evidence around video-based learning has led to a ‘proliferation of heuristics’ (Lodge et al., 2017, p. 2) in video production that remain largely untested, with the authors specifically referring to the ‘6-minute rule’ (where videos need to be less than 6 minutes) as an example of an untested claim (Lodge et al., 2017). Furthermore, of the studies of video-based learning using randomised or semi-randomised conditions, few have reached conclusive findings (Lodge et al., 2017). Giannakos, Chorianopoulos, & Chrisochoides (2015) state that learning via video-based lessons too often place students in the role of passive learners and that LA could provide insights into improvements of video-based teaching strategies.

Given that video-based learning is ubiquitous in online learning (Ou, Goel, Joyner, & Haynes, 2016) it is somewhat surprising that LA studies on video-based learning is a largely unexplored field (McGowan, Hanna, & Anderson (2016). Most LA research so far, has been in the areas of learner dashboards (Jivet, Scheffel, Specht, & Drachsler, 2018) and student recruitment (Siemens & Long, 2011). However, as discussed, understanding video-based learning is of vital importance to 21st Century learning and LA may be uniquely capable of offering solutions in this new learning paradigm.

Related Work

The literature review compared studies that used an LA method to investigate learner behaviour in video-based learning. Each study generally described the behaviours as engaged or disengaged rather than using the term 'active learning', but the behaviours examined (pausing, drop-outs, re-watching and note taking) are the same as those investigated by this study, and it could reasonably be argued that that these behaviours are consistent with active learning. In the discussion of these studies the term 'engagement' will be used instead of active learning when this was the term employed by the authors.

Of the few studies that have applied LA to video-based learning, a study by McGowan et al., (2016) evaluated the viewing behaviours of students in two courses on computer programming. The study used LA methods to identify student engagement with the videos and specifically sought to compare levels of engagement to traditional teaching methods like lectures (McGowan et al., 2016). The sample for the study was 80 students, across two modules, with all students also being surveyed about their viewing behaviours, which provided the study with additional qualitative data to add validity to its findings. The LA method employed analysed student data and presented it in the form of a visualisation where student behaviours like re-winding, skipping ahead and dropping out were revealed as 'peaks' and 'drop-offs' in a line graph. The study drew the inference that the smoother the graph or section of a graph the more engaged the students were with the video (McGowan et al., 2016).

The McGowan et al., (2016) study mapped student engagement to the content of the video, with one of their findings being that large amounts of text in the video had initial levels of high engagement but that if the text remained on screen for longer than 30 seconds, students would often skip forward (McGowan et al., 2016). Reviewing this finding against Mayer's (2008) cognitive principles for multimedia learning may offer an answer to this behaviour. The redundancy principle advises that text onscreen should not repeat (mirror) audio - making one essentially redundant (Mayer & Johnson, 2008). It could be argued that in the McGowan et al. (2016) study, students found the information they needed (through reading the slide) and made the decision that they did not need to listen to the lecturer essentially repeat verbatim what was on screen, and therefore they skipped ahead. McGowan et al., (2016) concluded that when students skip ahead in these sections it infers disengagement on the part of the student, however it may actually be indicative of engaged viewing i.e., the student is aware they have learned all they need from that section of the video, so move ahead to a new section. The McGowan et al., (2016) study found that students were most engaged during walkthrough explanations of coding, with students being more likely to repeat these sections and/or take notes.

A study by Kim et al., (2014) investigated 862 videos in a MOOC (Massive Open Online Course) on edX and used trace data harvested from the video viewing sessions to identify engagement. Similar to McGowan et al., (2016) a key finding from the study was that students watched more of the video in their first viewing session, while in secondary viewing sessions, the data revealed greater 'drop-outs' and more 're-watch' sessions (a section of the video being watched multiple times) (Kim et al., 2014). Again, as with McGowan et al., Kim et al (2014) interpreted this behaviour as inferring disengagement, and argued that in more engaging videos students may stay until later in the video (Kim et al., 2014). A further finding of the study was that there was a correlation between longer videos and higher drop-out rates, which the authors argue may be due to students' short attention span and/or feeling bored (Kim et al., 2014).

Evaluating the same behaviours, Lagerstrom et al., (2015), reached a different conclusion on student motivations behind the behaviour of ending a viewing session. As with the McGowan et al., (2016) and Kim et al., (2014) studies, the Lagerstrom et al., (2015) study investigated data around students' playing, stopping and re-watching, as well as jumping and speed change. However, this study identified an additional behaviour of returning to re-watch a video after an earlier session that was ended before the completion of the video (Lagerstrom et al., 2015). Interestingly, the authors interpreted the findings differently to McGowan et al., (2016) and Kim et al., (2014) in that they felt the results did not support the '6-minute rule' for video length in order to maintain student engagement (Lagerstrom et al., 2015). What they identified was that although there may be higher dropout rates when viewing individual sessions, students often returned to a video and that when the multiple viewing sessions were stitched together, the average percentage of a video watched by a student was close to 90% (Lagerstrom et al., 2015). Lagerstrom et al., (2015) considered many questions remain in video-based learning, stating that they can only infer student intent from the skips and pauses revealed in the data, and that more work was needed to understand underlying student intentions. This study employed a questionnaire to identify these underlying intentions.

A study by Dodson et al., (2018) was similar to previous work in that it used an LA method to explore student behaviours in video-based learning. However, it evaluated the behaviours against the ICAP framework for active learning. The ICAP framework for active learning is a listing and ordering of active learning into separate (sub)modes of engagement. According to Dodson et al. (2018), theirs was the first to apply the ICAP framework as the definition of active learning in video-based learning. They argued that, although there has been consensus that active learning improves the efficacy of video-based learning, there has been little theoretical understanding of what that active learning should look like (Dodson et al., 2018). Using the framework the authors identified the following types of video-based learning behaviour as being linked to active learning; browsing for general information, searching for specific information, triaging between learning objects (video and textbook), re-watching specific video content, and pausing to reflect on content or engage in another viewing behaviour (Dodson et al., 2018). The Dodson et al., (2018) study investigated undergraduates in a tertiary level science course and used what the authors described as a video player specifically designed for 'active viewing' (Dodson et al., 2018). This video player included a filmstrip which enabled students to re-wind or fast-forward using the strip as a thumbnail guide as well as including a transcript and a note-taking area (Dodson et al., 2018). The video-player was called ViDeX, and Dodson et al., (2018) argued that their study supports the contention that provided with a video player designed for active learning, learners will engage in additional active learning strategies to support their learning. The study investigated trace data logged by 28 students and as such the authors noted the relatively small sample size as a limitation (Dodson et al., 2018).

Dodson et al's. (2018) study is similar to this research project in that it used video-based trace data to identify active learning as described by the ICAP framework. It also supplements that data with an online questionnaire. However, where the Dodson et al., (2018) study differs is in the setting. The Dodson et al. study was set in an undergraduate course in a blended learning format where students interact with each other as well as lecturers/facilitators. The environment for this study, on the other hand, is a K-12 setting, which is an area identified as lacking these kind of studies (M. N. J. B. J. o. E. T. Giannakos, 2013). Another significant difference is that the HSC Study Lab is a wholly online environment with no facilitator support or guidance, whereas the Dodson et al., (2018) study was a blended environment with facilitator support. Finally, the video player (ViDeX) was specifically designed to allow additional behaviours such as video highlighting and note taking, with Dodson et al., (2018) arguing that provided the right tool learners will engage in active viewing behaviours. The HSC Study Lab video player is far more common to online environments, and allows only the usual viewing behaviours of playing, re-winding, fast forwarding, pausing, and stopping.

Two further studies have applied an LA method matched to the ICAP framework in order to identify active learning behaviours. Stepanek and Dorn (2017) developed a coding rubric based on the ICAP framework to investigate text artefacts created by students in an online learning environment. The authors gathered these texts and applied a machine learning technique to categorise them by levels of engagement (as coded against the ICAP framework) (Stepanek & Dorn, 2017). The study found the ICAP model they developed for identifying student engagement by analysis of text artefact was feasible and recommended further research (Stepanek, 2017). A second study by Marzouk, Rakovic, and Winne (2016) developed a tool called nStudy to gather data on when students highlight text, make notes, and/or search for information within the LMS. The authors used a data mining technique mapped against the ICAP framework to create data visualisations based on student engagement in the learning environment. It was hypothesized that provided a visualisation of their learning, the students would begin to self-monitor and thus develop meta-cognitive and self-regulated learning skills (Marzouk et al., 2016). The authors state the reason for choosing the ICAP framework (over other models of active learning) was that it describes overt learning behaviours that are identifiable with an LA method (Marzouk et al., 2016). Stepanek (2017), does not explicitly claim it as the reason for adoption over other models, but also states that the

ICAP framework focusses on visible behaviours rather than attempting to account for students' thoughts. Furthermore, it is not only that these behaviours are visible but they're mapped to underlying cognitive processes and that the claims of the study are supported by research (Marzouk et al., 2016; Stepanek, 2017). It is the contention of this study that this focus on visible behaviours makes the framework more applicable for studies that employ LA than other models of engagement.

Summary

This research project investigated evidence of active learning in an online learning environment (HSC Study Lab) that is wholly self-directed and with no instructor or facilitator support. This type of independent learning environment has been described as a growing phenomenon in online learning (Bowers & Kumar, 2015) and as such, this research project can be considered highly relevant to online educators. It is clear from the literature that there is a dearth of studies around LA and video-based learning, and that all of the studies have occurred only in the last few years, so it is argued that this is a relatively new field of research. Of the studies introduced, five explored the same type of data (engagement data associated with the features of video players) and of those only one used the ICAP framework as a model for defining active learning (in video-based learning). The literature review was unable to locate any studies that explored the use of learning analytics for identifying active learning in video-based learning for K-12 students. This study will add to the new and comparatively under-researched area of video-based learning by applying the methods to a new setting of an entirely online course for high-school students. Furthermore, research has been extended with the use of multiple data sets – specifically the use of a questionnaire that provided greater context and insight into the findings. Finally, the use of a multi-method approach helped to mitigate potential for misalignment between observed behaviour and underlying learner intent, as well as providing insight into student learning behaviours outside of (and thus not captured by) the online learning program.

Methodology

Research Problem and Questions

This research project used an LA method applied to video-based user data, within an online learning environment, to identify learner behaviours that have been linked to increased learning outcomes – namely, active learning as defined by the ICAP framework (Chi & Wylie, 2014). The project further supplemented those findings with the results of a questionnaire on the viewing habits of participants within the HSC Study Lab learning program. The questionnaire provided further insight into learner behaviour as well as additional information regarding underlying learner motivations when participating in those behaviours.

The research questions against which data for this this study were collected are:

1. To what extent do students participate in active learning behaviours in HSC Study Lab?
2. To what extent is learning analytics an effective tool for identifying patterns of behaviour associated with active learning in HSC Study Lab?

Methodology

In answering the first research question the study applied two data methods, an LA method and a questionnaire, both of which investigated patterns of behaviour identified as active learning. The second question was addressed in part by the LA method but primarily investigated through the application of the questionnaire. By extending the behaviours investigated and by mapping to underlying learner motivations the questionnaire provided insight into the question: To what extent is LA an effective tool for identifying patterns of behaviour associated with active learning in HSC Study Lab?

Theoretical Framework

The theoretical framework through which active learning is defined in this study is the ICAP framework (Chi & Wylie, 2014). The ICAP framework for active learning was selected as it provided a clear definition of active learning mapped to overt motoric behaviours. These motoric behaviours produce trace data in the video-based learning environment which are analysable using an LA method. In this study, the determining of active learning intent was enabled through the use of a questionnaire, which asked students to self-report on the intent behind their viewing behaviours. In order to illustrate the relationship between active learning behaviours and underlying student motivations the framework was expanded to include them. This study’s conceptual framework, identifying modes of engagement, associated behaviours and aligned motivations are summarised in Table 2.

Table 2

Conceptual framework summarising engagement mode, definition, aligned behaviour and underlying motivation/intention

Engagement mode	Definition	Observed behaviour	Leaner motivation/intention
Passive	is the lowest rung of the ICAP framework and is defined by the learner being oriented towards and receiving the information from the learning object or instructor, but not acting on or interacting with the learning object in anyway.	Playing the video	Basic, non-targeted information building
Active	is the second mode of engagement and can be identified by the learner acting on the learning object in a motoric or physical capacity.	Skipping forward or back within the video Re-watching sections of video Pausing the video Stopping the video	Information searching Reviewing, seeking clarification Reviewing, reflecting, seeking clarification Identifying that specific information needs have been met
Constructive	those behaviours that result in the production of additional outputs or products to the initial learning material, thus a characteristic of the mode would be that it is generative.	Taking notes while watching the video Explaining the video to a classmate Asking questions	Translating/extending understanding, linking concepts, Making inferences Translating/clarifying understanding Clarifying/extending understanding
Interactive	the highest level of engagement and like constructive, is generative, but with the additional requirement that the generative output was collaboratively created.	Collaborating with a peer or teacher to take notes or otherwise expand on the content of the video	Co-constructing, co-clarifying or co-extending understanding

Methods: Data and Analytics

Video Lesson from which Data was Captured

A video from HSC Study Lab (the online learning environment) was selected and an aggregate of second by second user interaction data was analysed. A decision was made to use a video from the Y12 biology program as Y12 biology contained the greatest number of users (of the three science programs covered) and as such, contained a larger sample size of the total student cohort. Furthermore, the video chosen was one from the second year of the program (the program covers years 11 and 12) as it is argued that Y12 users would be more familiar with the format and potentially have more established video viewing habits. The number of total plays for the video at the time of analysis, was 870. The title of the video investigated was 'Innate and adaptive immune system' and was an animated video on the human immune system. It is beyond the scope of this study to map student engagement to specific content or themes within the video. However, it is understood that such an approach would add an extra dimension to understanding student learning behaviours in video-based learning. This has been the approach of related studies such as McGowan et al., (2016), however as discussed it presents a challenge for large scale LA based studies as it requires viewing of each video by the researcher. Future research will investigate the possibility of mapping learner engagement data to in-video content.

Data Capture Method

The first stage of a learning analytics research project is the capture of data (Pardo, 2014). The main form of data captured was user actions while watching videos. HSC Study Lab uses an external hosting service for streaming its videos, and this service (Sprout), allows for the capture and visualisation of data associated with playing, pausing, skipping and re-watching the video (Figure 1). Sprout provides analytics through visualisations but does not allow easy access to the underlying dataset, which resulted in the necessity of second by second transcription to transpose the data into a usable form. The video hosting platform keeps a record of how much of a video the user watches, whether they skip sections, as well as if they re-watch (and how often they re-watch) sections of a video (within one viewing session). If the user exits the video completely (closes the browser) it is considered the end of the viewing session and the percentage of the video watched is recorded. If the student returns to the video later this is treated as a new session.

As can be seen in Figure 1 the video service provides visualisations of the data representing the viewing session of individual users. The time stamp at the bottom of the bar indicates the run-time of the video, the percentage figure at the end of the bar informs of the overall percentage (not necessarily sequential) of the video watched. Finally, the colour of the bar indicates what sections were watched, and how often. The hue of the colour bands within the video indicates whether a section was re-watched, with the colour changing in intensity and colour (darker green and then yellows and reds) depending on the number of times that section of the video was watched. Figure 2 illustrates a student viewing session where sections of the video were watched multiple times. The colours within the band and the number of times that section is watched are given in Table 3.

Figure 1

Visualisation of data on the viewing session of individual users

Figure 2

Bars of colour indicating number of times section of video is watched

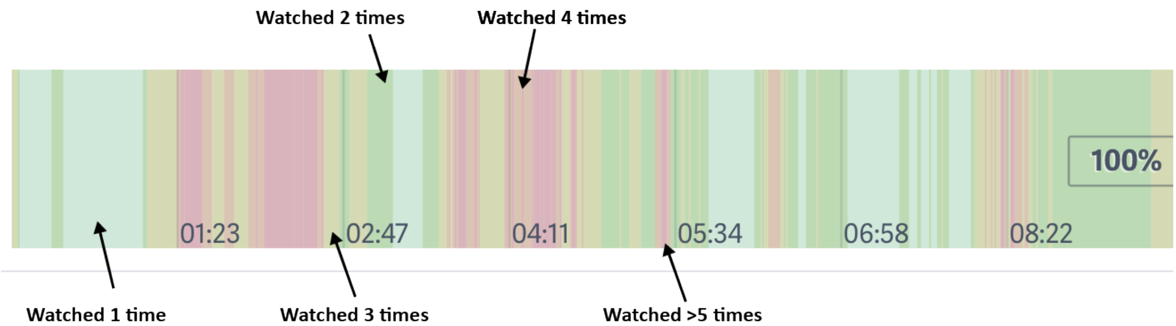


Table 3

Colour of band within session visualisation and number of times that section of video was watched

Colour of video band	Number of times section watched
Light green	watched 1 time
Dark green	watched 2 times
Yellow green	watched 3 times
Light pink	watched 4 times
Dark pink/red	watched 5 times (or more)

Data Coding (Learning Analytics)

Data harvested from the digital video player is mapped against student behaviours in the learning environment. Once the data is mapped to student behaviours it is possible to use the framework to identify the mode of active learning. For example, re-watching sections of a video, along with pausing, skipping and even stopping are behaviours consistent with Active engagement (a sub-mode of active learning). Conversely, watching a video without otherwise acting upon it, is Passive engagement (Chi & Wylie, 2014). This is detailed in Table 4.

Table 4

Map of viewing behaviour against engagement level

Viewing behaviour	Engagement level
Viewing a section of video multiple times	Active
Viewing a section of video once	Passive

In order to further illustrate the data visualisations and their meaning, two examples are presented. In Figure 3 the unchanging light green colour and the ‘100%’ reading at the end of the bar indicate that this viewer (during this session) watched the video in its entirety from start to finish without skipping ahead, or re-watching any sections within it, which would infer Passive engagement. Figure 4, on the other hand, reveals a viewing session where there is a consistent pattern of re-watching sections of the video multiple times. Figure 4 also reveals that the user only watched 85% of the video before dropping out. With consideration against the framework this pattern of viewing would indicate Active engagement.

Figure 3

Visualisation of data where 100% of video was watched and no sections re-watched

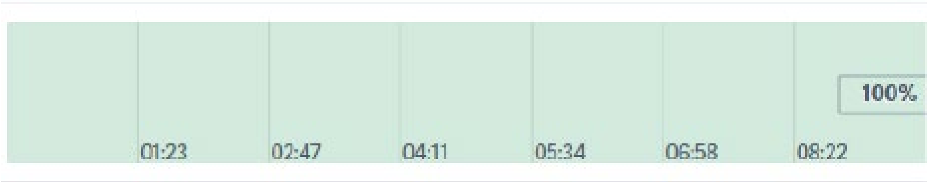
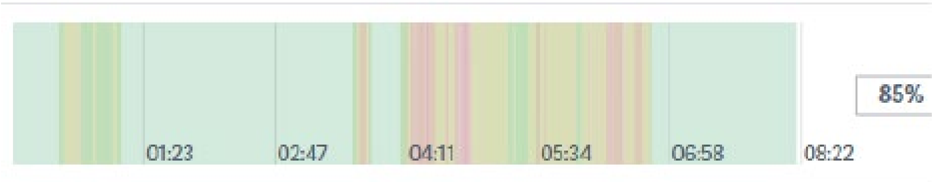


Figure 4

Visualisation of 1 session, 85% of video watched, colour bands indicate re-watched sections



Each video also has a visualisation of the aggregate data associated with all viewers and viewing sessions, as detailed in Figure 5. The figure is a graph with a time stamp along the X axis and a percentage representing the total number of views on the Y axis. Total number of views is not the same as total number of viewers, as viewers may re-watch sections multiple times. Peaks in the graph are caused by students re-watching sections of the video while dips are caused by students dropping out or skipping ahead. This movement in the viewing percentage can be mapped against modes of engagement as defined by the ICAP framework. The relationship between the movement of the graph, the behaviour and the mode of engagement is detailed in Table 5.

Figure 5

Visualisation of the aggregate viewer engagement with the video

Table 5

The relationship between line graph, indicated behaviour, and aligned engagement mode

Line graph movement	Viewing behaviour	Mode of engagement
Straight (unchanged)	Viewing a section of video once	Passive
Peaks (line moves up)	Viewing a section of video multiple times	Active

As part of the data preparation process the viewing percentages were taken from the video, second by second, and entered into an excel spreadsheet, which created two data columns, one for time and the other for percentage of viewership. The data associated with time was converted to seconds (instead of seconds and minutes) to support analysis. Table 6 contains a sub-section of data illustrating the relationship between time and viewership.

Table 6

Percentage of viewership against time in the data

Time (s)	Viewership (%)
32	81%
33	81%
34	81%
36	79%
37	78%
38	78%
39	77%
40	75%
41	74%
42	75%

Data Analysis and Definition of a Peak

There is a general decrease in viewership across the length of the video caused by the user dropout rate (the rate at which viewers leave the viewing session), which tends to mask the significance of the peaks. For example, it is difficult to identify an increase in viewership when each successive peak is likely to be less than total viewership at the start (or earlier) in the video. Therefore, a working definition of a peak that took into consideration this trend was required. When data was transposed to an excel worksheet and converted to seconds it was possible to apply a simple formula to account for the general decrease caused by the dropout rate. Given the data available an ad-hoc analysis technique was considered most suitable (Pardo, 2014). The following statistical formula was applied, adjusted and ultimately adopted for use in this study.

The formula selected a data point $\%t_2$ (the data point for which a reading is sought) and subtracted a comparison data point $\%t_1$ (a data point a number of seconds earlier in the video) which gave a final percentage reading ($d_{\%}/d_t$), which was either greater or lesser than the data point it was compared with. A 'trigger' was then created that would call a 'peak' (P) depending on how great that difference in viewing percentage was (provided it was also always a positive difference). The trigger was necessary in order to account for noise and to smooth out the data i.e., there are near constant changes in viewing percentages so a measure for meaningful change was required. When deciding the number of seconds over which a peak was called, consideration was given as to what would be a meaningful ‘chunk’ or section of video to be re-watched. As this is a new area of study there was little guidance within the research as to what could be considered a meaningful or significant chunk of video for re-watching. This was ultimately decided by the width of the resulting peaks, but further analysis against multiple videos is required to add validity to this method. For example, initially a timeframe of 10 seconds was selected and the outcome this produced was more peaks but the average width (duration of viewing) for each peak was 6.4 seconds. Increasing this timeframe to 20 seconds the analysis resulted in fewer peaks but the average duration (or width) of each peak increased to 15.83 seconds. This research project selected the longer timeframe of 20 seconds but as stated, further comparison across multiple videos is required to establish a baseline for significant events. Therefore, for this study the comparison point was 20 seconds earlier in the video so the timeframe over which peaks were called were 20 second blocks. The trigger for calling a peak was a 5% increase in viewership, which means that if there was a 5 percent increase in viewership over any 20 second timeframe a ‘peak’ was called. The formula employed was as follows:

$d_{\%}/d_t = \%t_2 - \%t_1/t_2 - t_1$, where $t_2 - t_1 = 20_s$

Under these conditions a ‘peak’ is defined where: $d_{\%}/d_t >= 0.5_{\%}\ 5^{-1}$

Methods: Questionnaire

A web-based questionnaire was developed and sent out to students enrolled in year 12 of the HSC Study Lab program. Year 12 students were selected as they were likely to have had more experience in the program (it is restricted to years 11 and 12) and as such would be more familiar with the tools and style of videos. Year 12 corresponds to the sample video selected (Innate and Adaptive Immune System) and it is highly likely, considering Biology is the most popular of the three programs offered, that some of the participants of the questionnaire represent a sample of the data associated with the video. The questionnaire was emailed to Y12 HSC Study Lab students, with 106 responses being received.

The purpose of the questionnaire was, in the first instance, to validate the findings of the LA method as well as evaluate inferences that learner intentions behind identified behaviours conformed with active learning. This is because in order for learner activity to represent active learning, there must be a corresponding intent on the part of the learner. This is a common feature of active learning models with Bonwell and Eison (1991) stating that active learning behaviours must be meaningful i.e., they must be linked with 'higher-order' cognitive processes like 'analysis' and 'synthesis' in order to be considered active. This requirement is also found in Scardamalia and Bereiter's (2006) concept of 'intentionality' in active learning, while it is called 'focussed attention' in the ICAP framework (Chi & Wylie, 2014). An additional role of the questionnaire was to identify non-program-based engagement with the video-based lessons, for example note taking or discussing the videos with classmates and teachers. A web-based questionnaire was developed and sent out to students enrolled in year 12 of the HSC Study Lab program in July 2019. The questionnaire is a common and useful instrument for collecting information that is structured, numerical and comparatively straightforward to analyse (Cohen, 2007) and given that this study investigated large numbers of users, a structured and quantitative questionnaire with nominal scale questions was considered suitable. Furthermore, as it is difficult to make sense of the detail provided by data at the micro level (Coffrin et al., 2014) - the second method improved understanding. As most of the questions in the questionnaire are nominal in design they were easily coded and processed using the web-based survey tool Qualtrics.

Data Coding (Questionnaire)

The analysis of the video player trace data allowed the researcher to identify patterns of behaviour that could be categorised as passive or active (including all sub modes as defined by the framework). The questionnaire was designed to augment these findings by asking participants to self-report on the behaviour. For example, as it is possible within the trace data to identify re-watching of sections of the video, a question specifically asked participants to confirm that they participate in that behaviour. If the participants responded in the affirmative, then there is support from the questionnaire that the patterns of behaviour identified in the trace data are an accurate reflection of learner behaviours. Furthermore, as discussed it is not enough that the behaviours conform to active learning, the intention or student motivation behind the behaviours also need to align with the observed (or self-reported) behaviour. Additional questions were designed to elicit responses that illuminate the motivations behind the behaviour. The questions could be categorised as belonging to either the 'environment', 'observable behaviour' or 'motivation/intention'. The responses to the categories 'observable behaviour and 'motivation/intention' could be coded against the framework and mapped to the specific sub modes of engagement within active learning. This relationship is illustrated in Table 7.

Table 7

Category and coding of response against ICAP framework

Question	Observable Behaviour	Learner Motivation/Intention	Response
Q5	Stopping the video		Active
Q6		Found what I needed	Active

Participant responses to follow up questions (questions associated with motivation/intention) were also coded against the active learning sub modes. A univariate analysis of the results was applied with students self-reporting as participating in the behaviour (or not) along with a general frequency. By using the questionnaire, the study could provide further insight into student viewing behaviours within HSC Study Lab, including whether their underlying (and invisible to the LA method) intentions also align to active learning as defined by the framework.

Results

Data Analysis

Individual Sessions

The analysis of the visualisations associated with viewing sessions revealed variation in viewing habits among individual users. Some users began a video and watched it without interruption, and until completion. Figure 6 illustrates this as an unchanging light green colour in the data visualisation, indicating a reading of 100% viewed. Conversely, Figure 7 illustrates a much more engaged viewer. The variation in the colour band throughout the timestamp indicates a pattern of behaviour where several scenes were watched multiple times.

Figure 6

100% of video watched with no interaction

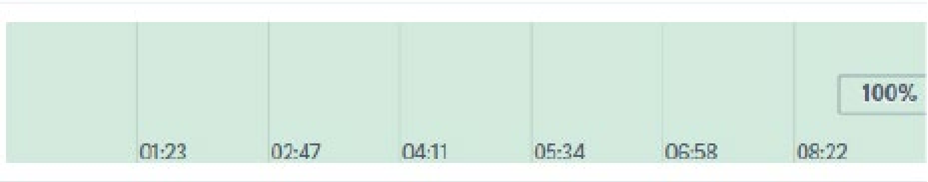
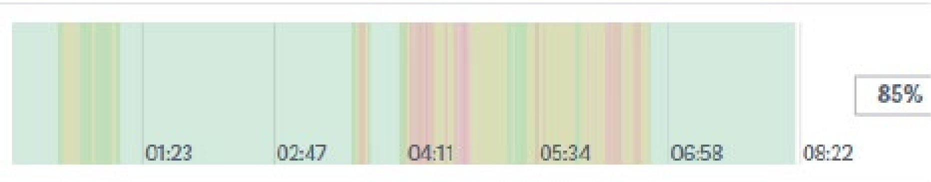


Figure 7

85% of video watched with colour bands clearly indicating a pattern of re-watching



By linking separate viewing sessions with a shared IP address, it was possible to identify students returning to a video and completing it over multiple viewing sessions. In Figure 8 a viewer (on the 01/08/19) starts a viewing session, re-watches sections earlier on in the video, then re-watches sections from approximately 3 min to 6 min multiple times before leaving the video around the 8-minute mark. Figure 9 suggests a user with the same IP address began a new viewing session of the same video, but 20 days later. In the second session (Figure 9) the user appears to skip over the first 3 minutes of video which was the same three minutes they showed limited engagement with in the first session. Then, there is little to no re-watching of the video, and this time the video is completed. This may indicate that the student watched this video, finding the content between the 3- and 6-min mark of most interest or relevance, and then revisited 20 days later for a refresher, this time jumping directly to the section that they found of most interest the first time. The final pairing of data visualisations (Figure 10 & 11) also indicates a viewer returning to and completing a video over multiple viewing sessions. In Figure 10 a user (on 18/07/2019) starts the video and watches until around the 6-minute mark before dropping out. In this session, they appear very active as they re-watch multiple sections and even skip over some sections. In Figure 11, it is revealed that the next day (19/07/2019) a user with the same IP address re-enters the video, skips ahead until they reach approximately when the previous session was ended, and then watch the video until completion, again re-watching a large section and smaller sections multiple times.

Figure 8

Pairing of two viewing session with shared IP address

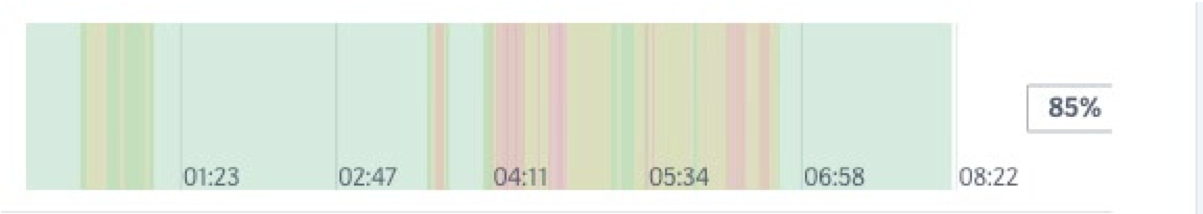


Figure 9

Pairing of two viewing session with shared IP address

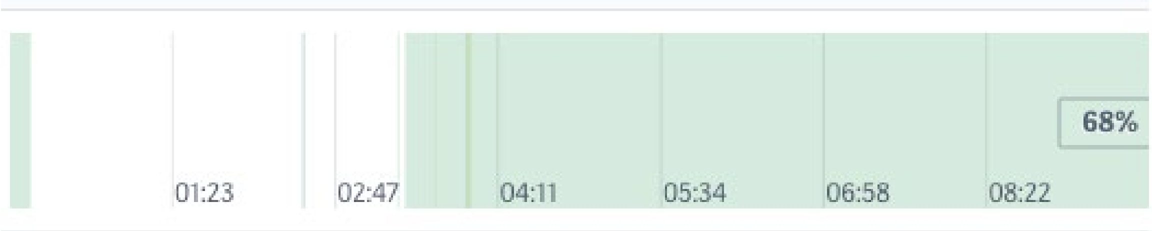


Figure 10

Pairing of two viewing session with shared IP address

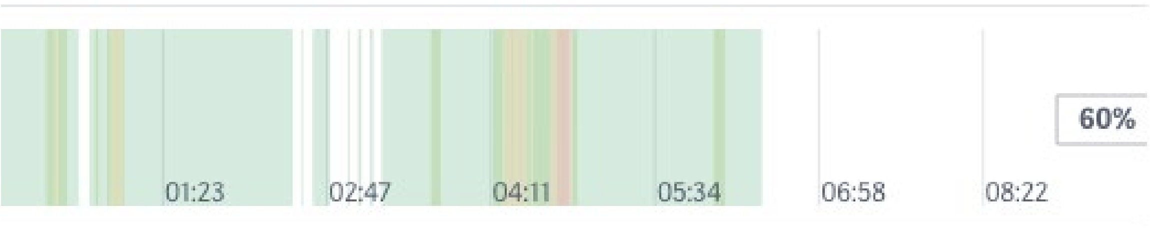
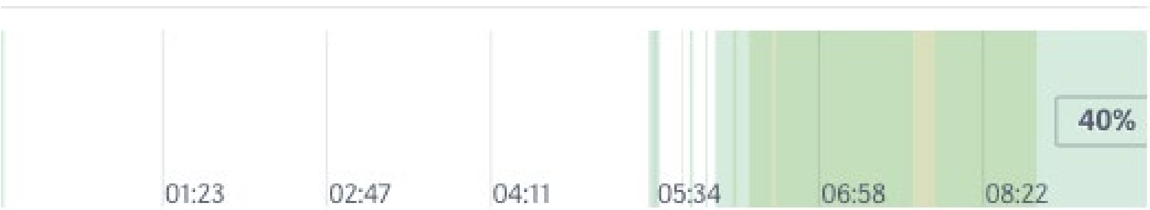


Figure 11

Pairing of two viewing session with shared IP address

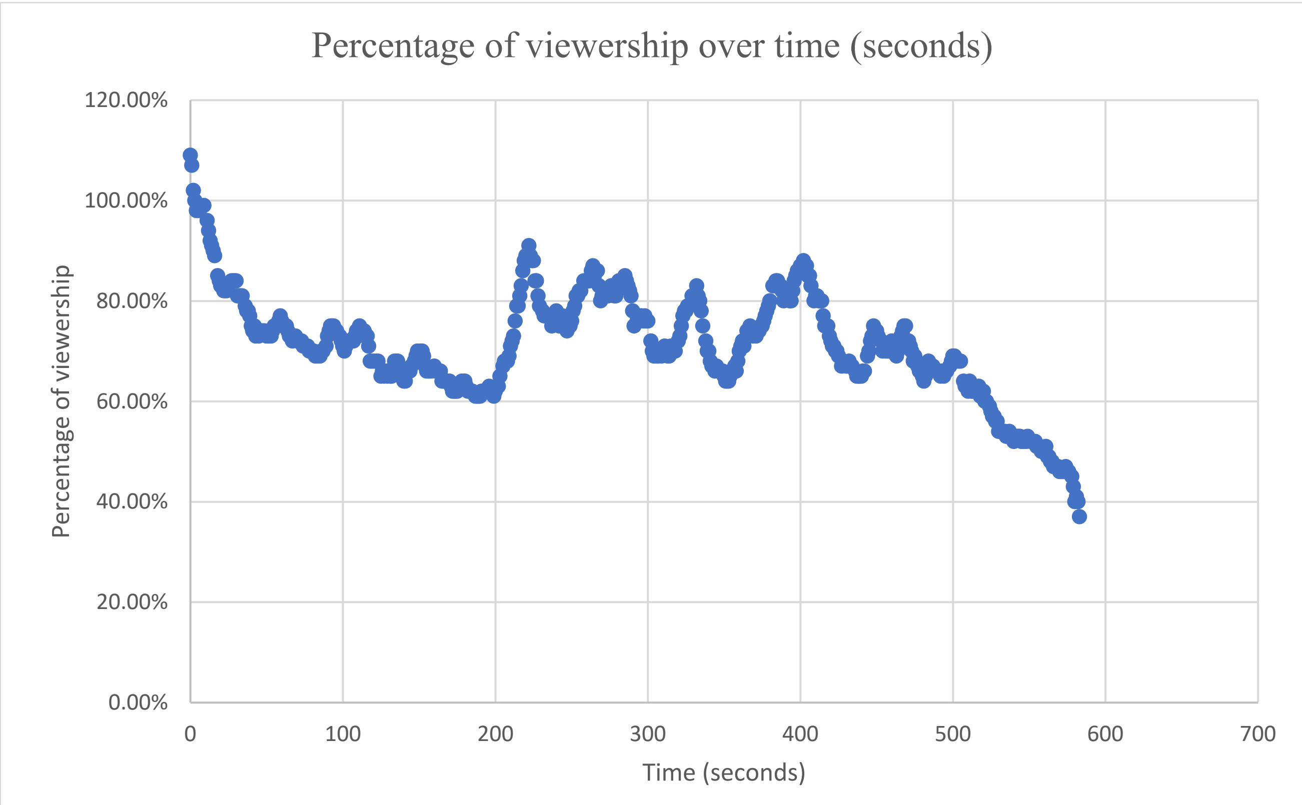


Aggregate Data

Aggregate data was harvested from a Y12 biology video identifying the viewing behaviours of students. The video had been viewed 870 times at the time of data harvesting, which provides a large sample size and adds to the reliability of the patterns of engagement visible in the individual sessions. Once the data was entered into an excel spreadsheet the graph illustrated in Figure 8 was created. There is a large initial percentage of viewership (of over 100%), a relatively even (and steep) drop off until around the 200 second mark, and then a series of peaks and troughs until the 500 second mark and then another steep drop off (ending with below 50% viewership). What is also illustrated is that this drop off in viewership is not even- that there are a series of peaks in viewership starting around the 3-minute mark and fluctuating until around the 7-minute mark, before it starts to more uniformly drop off. These peaks are caused by an aggregate of all the re-watched sections of the video by the individual users. So, although there may be a general drop off in viewership caused by the dropout rate this is countered in the data by the peaks caused by users re-watching specific sections of video multiple times.

Figure 12

Percentage of viewership over time



A formula was applied to the data to reduce noise and acquire an indication of the number of peaks across the timescale of the video. The formula used was:

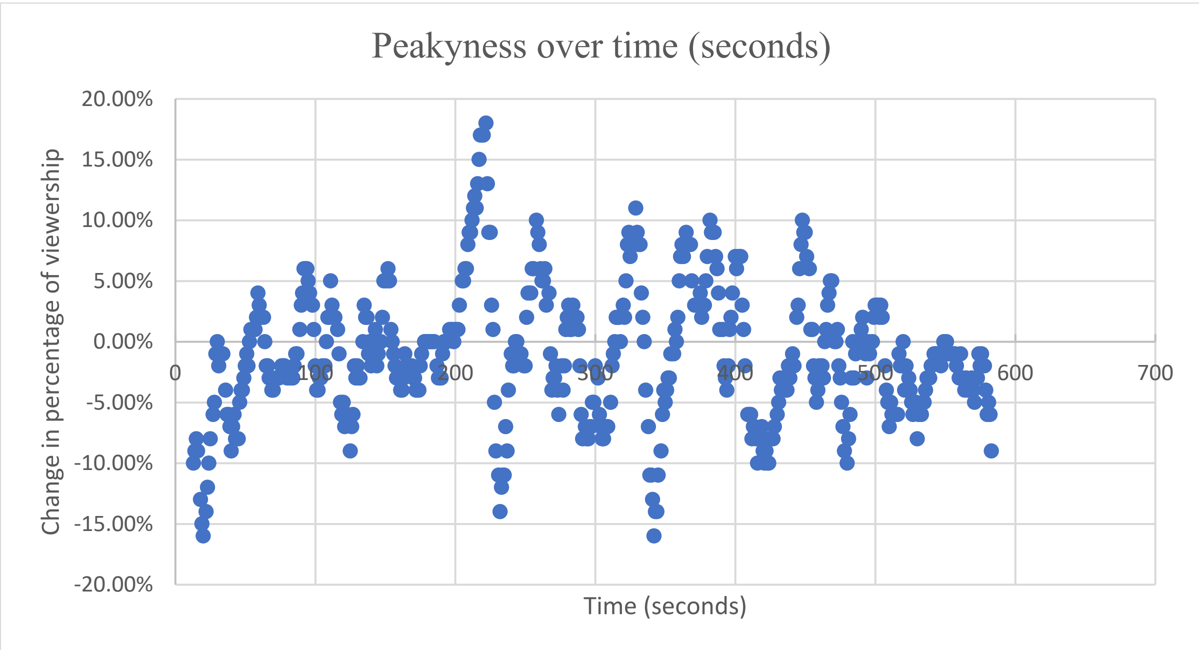
$d_{\%}/d_t = \%t_2 - \%t_1/t_2 - t_1$, where $t_2 - t_1 = 20_s$

Under these conditions a ‘peak’ is defined where: $d_{\%}/d_t > = 0.5_{\%}\ 5^{-1}$

This formula allowed for a comparison of ‘peakyness’ (height and size of peak) between the varying data points. The data was then re-graphed, and the following visualisation (Figure 13) was created. Along with the visualisation the formula revealed a total number of 6 peaks at 152, 222, 263, 332, 382, and 449 with an average increase in height over the 20 second timeframe of 9% and an average width of 15.83 seconds. Both the timeframe for comparison and the value of the trigger require further comparison with additional videos in order to improve reliability of baselines for significant events.

Figure 13

Graph revealing peakiness of data over time



Questionnaire Results

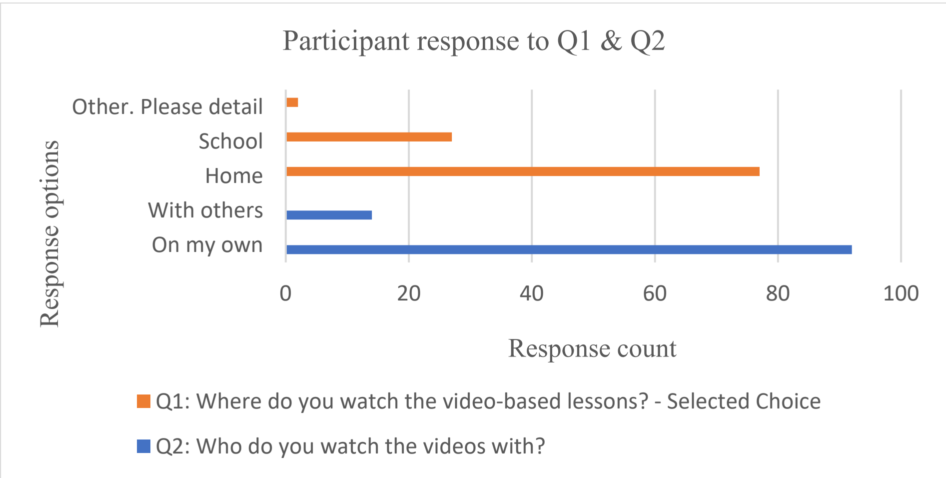
A questionnaire was emailed to Y12 HSC Study Lab students, with 106 responses being received. The questions have been assigned and data coded under 3 categories; 1) the Environment, 2) Observable behaviour, 3) Underlying motivation/intention'.

The Environment

To establish that HSC Study Lab was in fact an independent learning environment with little to no teacher support, questions 1 and 2 (Figure 14) asked students where they watch the videos and with whom.

Figure 14

Participant responses to question 1 and 2



Observable Behaviour

The second category of questions dealt with viewer behaviour when viewing videos. These behaviours can be mapped to modes of engagement in the framework as detailed in the methodology section, and students self-reporting as participating in these behaviours indicate active learning. Responses to questions 5, 6 and 7 are represented in Figure 15. The behaviour addressed by question 5 can indicate Active engagement or Passive engagement depending on the motivation behind it, therefore question 6 sought clarification as to what that motivation was, and the results were coded and summarised in Table 8.

Figure 15

Participant results for question 5, 6 and 7

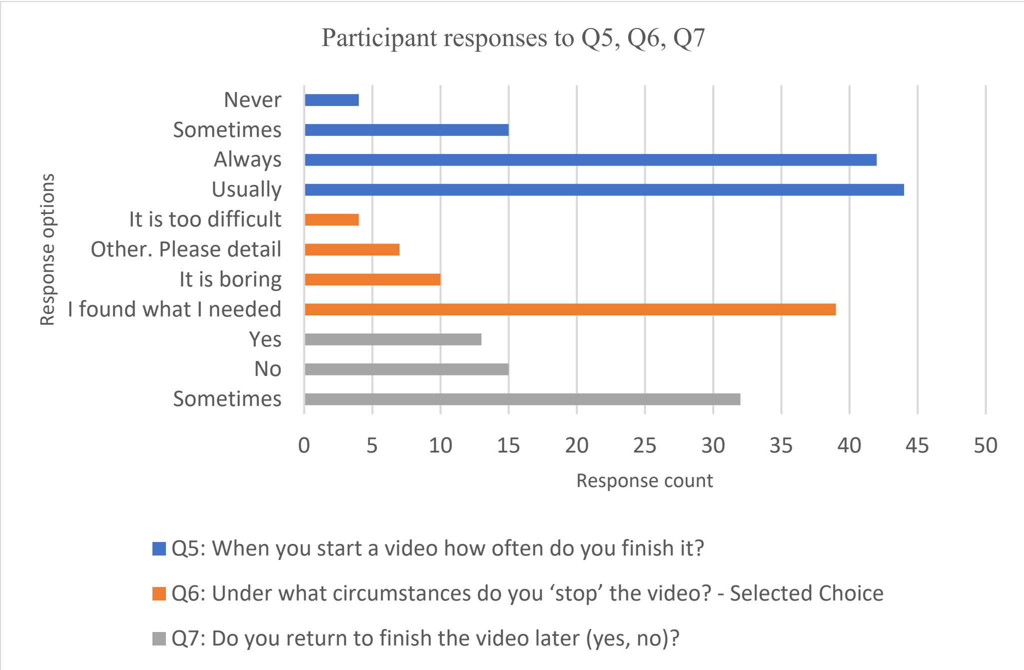


Table 8

Q6 response count and the coding against the framework

Response	Coding	Count
It is too difficult	Passive	4
It is boring	Passive	10
I found what I needed	Active	39
Other	N/A	7

The participants who responded 'Other' were asked to provide a detailed response. These included responses like "I find out it isn't relevant" and they too could be coded against the ICAP framework as summarised in Table 9.

Table 9

Responses to 'Other' in question Q6 coded against the framework

Coding of response	Count
Passive	0
Active	6
Constructive	1

Questions 8, 10, 11, and 13 focused on behaviours that allowed direct coding to modes of active learning as identified by the framework. Furthermore, as the questions allowed for basic yes/no responses the results could be aggregated. The specific questions are given in Table 10 and the student responses are provided in Figure 16. Question 9 was a follow up to question 8 and asked participants to identify where they take notes. The results are given in Figure 17. For those who responded 'Other' to question 9, additional details were sought and are summarised in Figure 18.

Table 10

The specific text of questions 8, 10, 11, and 13

Question	Coding against framework
Q8: Do you take notes while watching the video (yes, no, sometimes)?	Constructive
Q10: Do you pause the video to take notes (yes, no)?	Active
Q11: Do you discuss the content of the video lessons with another person (yes, no)?	Interactive
Q13: While watching a video do you skip back to 're-watch' parts of it?	Active

Figure 16

Questions 8, 10, 11, 13 coded to engagement level and participant responses

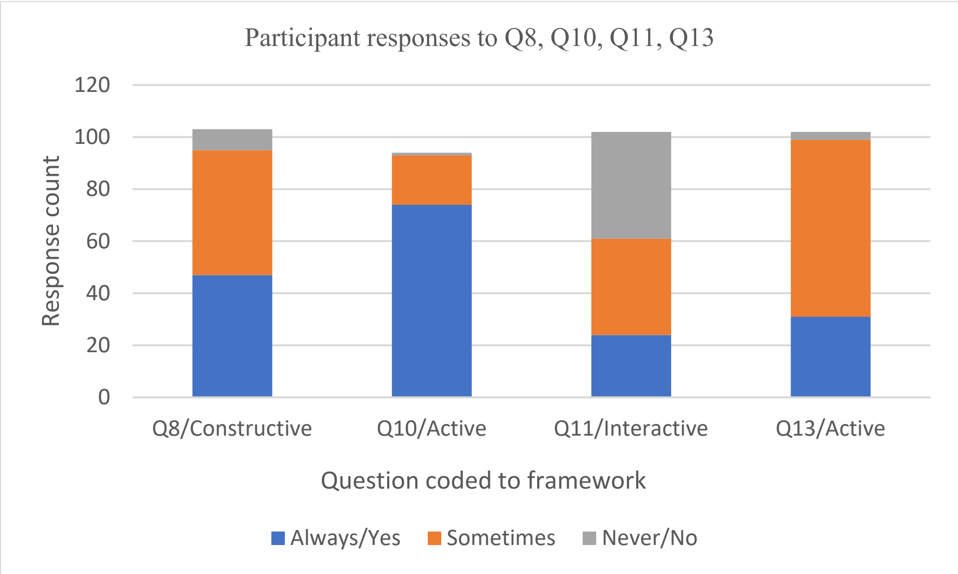


Figure 17

Responses to question 9: Where do you take notes?

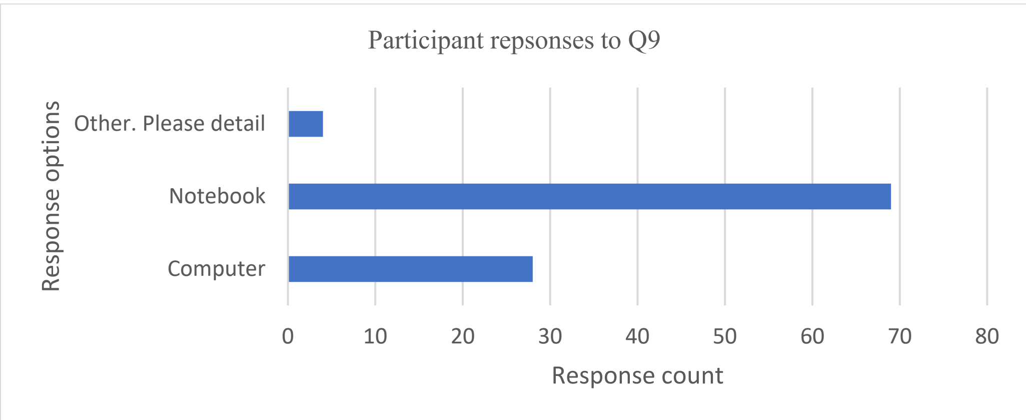
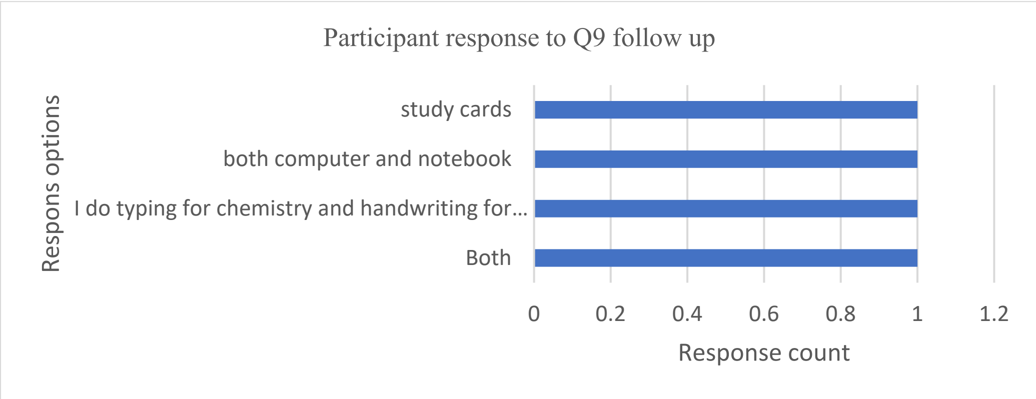


Figure 18

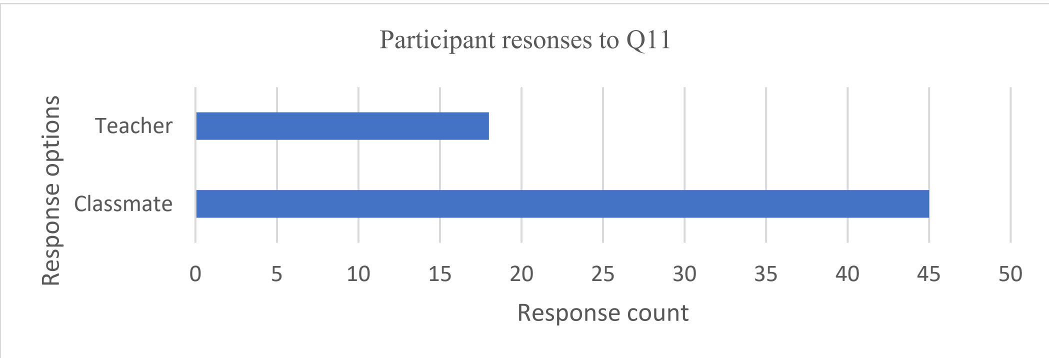
Results to question 9 follow up: Other please detail.



Question 12 was a follow up to question 11 and asked participants who they discussed the videos with. The results are given in Figure 19.

Figure 19

Results to question 12: Who do you discuss the videos with?



Question 17 asked participants to identify *'what viewing strategies they felt applied to them'*. The results to the question could be coded against the framework and are summarised in Table 12. An aggregate of the coded results against the ICAP framework is provided in Table 13.

Table 12

Participant responses and coding to question 17

Q17	Answer	Coding of response	Count
Summary of Q17: When watching a video which strategy/s applies best to you	Watch everything at once	Passive	26
	Take notes while watching	Constructive	77
	Watch the video with a classmate	Interactive	8
	Search the video for 'important points'	Active	43
	Other. Please detail		6
Q17 Other. Please detail	Answer	Coding of response	Count
	I could never finish one	Passive	1
	Watch the video, person says a concept im not familiar with and if they explain it but i still dont understand it, i pause the video and research it online until it understand it, then i resume the video and repeat	Active	1
	do questions at the same time	Active	1
	the first time i watch it I watch it without stopping, then i watch it a second time to get notes down	Constructive	1
	watch once and take notes the second	Constructive	1
	watch video through once, then play again with pen in hand	Constructive	1

Table 13

Responses to question 17 when coded against the framework

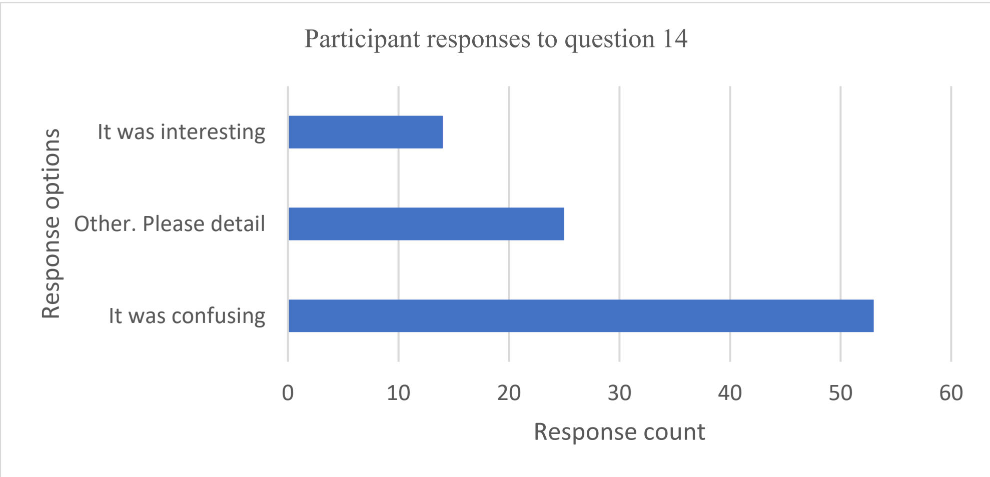
Coding of response	Count
Passive	27
Constructive	80
Interactive	8
Active	45

Underlying Motivation/Intention

The third category of question was ‘Underlying motivation/intention’ and were questions that asked participants to reveal the reasons behind their self-reported behaviour. Question 13 asked students whether they re-watch parts of a video. For those who answered that they did re-watch videos, a follow up question was asked in order to identify the motivation behind the behaviour. The responses to this question were coded against the framework and the results are presented in Figure 20.

Figure 20

Results to question 14: Why do you re-watch the videos?



Of those who answered ‘Other.’ additional detail was sought. As the answer took a qualitative form, the results were analysed, themed, and coded against the framework. Responses included “For clarity, or to ensure I understand.”, which could be coded to Active engagement. Another example included “So I can write what they said.”, which is coded to Constructive engagement. Table 14 provides the total number of responses coded against the ICAP framework.

Table 14

Responses to the follow up question to Q14 coded against the framework

Coding of response	Count	Percentage
Active	15	60%
Constructive	10	40%

Summary of Results

In the environment section of the questionnaire, question 1 and 2 (Figure 14) revealed that participants generally watched the videos independently with 72.64% of participants self-reporting that they watched the videos at home and 86.8% that they watched the videos alone. This finding confirmed that HSC Study Lab was an independent learning environment and it was therefore possible to explore the question 'to what extent do students participate in active learning behaviours in HSC Study Lab?'

Responding to the first question the data analysis supports the contention that LA can identify patterns of behaviour associated with active learning. In the analysis of individual viewing sessions, Figure 6 revealed patterns of behaviour that align to Passive engagement while Figure 7 revealed patterns of behaviour that align to Active engagement. In the aggregate analysis of data, the peaks in the visualisation (Figure 12) revealed multiple users re-watching specific sections of the video, which is evidence of Active engagement. When the data analysis formula was applied it revealed a series of peaks within the data. As could be seen in Figure 13 the largest peak comes at the 222nd second and is an increase of 28% over the given timeframe, further peaks occur at 152, 263, 332, 382, and the 449th second mark. The cluster and size of these peaks infer that there was content within those sections of the video that students felt particularly engaged with, whether that was due to interest, confusion or difficulty of the subject matter could not be identified by the LA method alone. What is clear, however, is that there is non-random student engagement with the video in the form of re-watching specific sections. This data analysis reveals patterns of behaviour that fit Active engagement as defined by the ICAP framework.

The results of the questionnaire supported the findings of the LA method as students self-reported that they did participate in the behaviours identified, and further, that they did so for reasons consistent with active learning motivations. For example, the largest reason given for leaving a video was that the student had found what they needed (Table 8), which aligns with an Active motivation, that students paused a video in order to take notes - a Constructive engagement motivation (Figure 16), and that they re-watched sections of a video to increase clarity, which aligns with Active engagement (Table 14). The mean of results for questions 8, 10, and 13 revealed 96% of respondents always or sometimes participated in active learning behaviours in video-based learning. Question 13 (Figure 16) addressed the behaviour of re-watching, which was also the behaviour identified by the LA method. The results of the questionnaire align with the findings of the LA method with participants responding that they always (30.4%) or sometimes (66.7%) participate in the behaviour. Finally, in question 17, participants were asked to indicate what viewing strategies they participated in. The results support the contention that students participate in active learning behaviours with 72.6% indicating that they participate in note taking. Further findings from the viewing strategies question were 40.6% of participants responding that they 'search the video for important points' (Active engagement) as well as a small percentage of participants (7.5%) indicating that they 'watch the video with a classmate'. Searching can also be observed in the trace data and appeared as a dip in engagement, essentially the opposite of a peak.

When considering the second question, 'to what extent LA is an effective tool for identifying patterns of behaviour associated with active learning' it was important to evaluate what alternative modes of active learning the questionnaire revealed that were not identifiable by the LA method. In Question 8 the students were asked 'do you take notes while watching the video?' with a total of 92% of respondents either answering 'yes' or 'sometimes'. This is a behaviour that aligns with the Constructive engagement mode of active learning. Constructive engagement behaviours are those that result in the production of additional outputs or products to the initial learning material (Chi & Wylie, 2014) and are considered a level higher than Active in the ICAP framework. As such, the mode is also associated with higher level cognitive engagement and learning outcomes (Chi & Wylie, 2014). What can be inferred from this result is that there were student behaviours and motivations in the video-based learning that were beyond the LA method (alone) to identify. This is due to the behaviour occurring using learning tools (notebook or computer) that sit outside of the learning environment, and therefore do not create analysable trace data.

The results also revealed that there was a misalignment between observed behaviour and student intention as predicted by the ICAP framework. Question 10 (Figure 17) also related to note-taking and asked participants 'Do you pause the video to take notes (yes, no)? All but one of the participants self-reported that they do pause to take notes while watching the video. In an analysis of the efficacy of the LA method, the ICAP framework was useful as it mapped observed behaviours to underlying motivations. Pausing aligns with the Active mode of engagement in ICAP, and is associated with motivations like 'reviewing', 'reflecting', and 'seeking clarification', while note taking is a Constructive behaviour, therefore there is a misalignment. It was through the use of the questionnaire that it was possible to have students confirm that their motivations matched that predicted by the framework. This behaviour of pausing to take notes was also not identifiable by the LA method as it did not produce distinguishable trace data. Question 11 (Figure 16) asked participants whether they discussed the content of the videos with others, with 58.8% of participants reporting that they do. These conversations, however, also sit outside of the digital environment and are therefore not identifiable by the LA method.

Where there was a high level of alignment between observed behaviour and underlying motivation was in the responses to question 13 and its follow up question 14 (Figure 16 and Table 14). Re-watching is a behaviour that is mapped to Active engagement in the ICAP framework and the associated motivations aligned to this behaviour are reviewing (improving or solidifying information) and seeking clarification, both of which are clearly visible in the responses from the questionnaire. A total percentage of 72.8% of participants responded that they re-watched sections of a video because it was either confusing (57.6%) or interesting (15.2%), strongly aligning with motivations consistent with an Active mode of engagement. Of those who answered ‘other’, a coding of the responses to the follow up ‘Please detail’ question revealed further Active motivations, as well as some Constructive. Furthermore, no participants reported ‘video error’ or other technical reasons that were unrelated to active learning viewing motivations as a reason for re-watching a section of the video. According to the ICAP framework there are improved learning outcomes and positive knowledge change processes associated with this pattern of behaviour (Chi & Wylie, 2014).

Discussion

The findings of the research project had relevance to earlier LA and video-based learning research and the following is a discussion of the major themes.

Re-Watching of Videos

The main behaviour investigated by both the LA method and the questionnaire was re-watching of sections of the video. The ICAP framework (Chi & Wylie, 2014) and the Giannakos et al., (2015) study identified improved learning outcomes associated with re-watching, so it is encouraging that 97% of participants within this study responded that they engage in such behaviours at least some of the time (Figure 16). When analysing the stated motivations behind the behaviour (Table 14), the majority of responses adhered to a theme of seeking clarity or improved understanding. This finding may infer that there are problems with that section of the video and that the subject content is not being clearly explained and/or is beyond the level of the student. This information can inform course facilitators and instructional designers and lead to improvements in the design or presentation of the video content. What is important to note, however, is that regardless of what it may infer about the specific content of the video, being able to recognise what is unclear or puzzling is an essential component of active knowledge building (Scardamalia & Bereiter, 1993). Furthermore, participants in the questionnaire reported at least some of the time this behaviour was due to ‘information seeking’, which is a highly engaged learning strategy associated with self-regulated learning as described by Zimmerman (1990) and Winne (2018).

As important as the specific pedagogical strategies and content design employed by the videos are, viewing strategies could be an overlooked component of the learning process. The fact that there is variation in the viewing strategies of the individual learners in this study infers that the content of the video itself may not be the only or even main motivator for the behaviours. The students do not appear to be passively viewing video content and where previously behaviours like skipping forward or dropping out may have been identified as revealing a lack of engagement (Kim et al., 2014; McGowan et al., 2016), it is possible they are evidence of active learning behaviours like information seeking. It is argued that active viewing behaviours, regardless of the specific video content, have benefits to students, and that LA identifying such behaviours has value in and of itself. This value proposition of LA is similar to that of the study by Marzouk et al., (2016) who focussed on student learning behaviours within online learning environments rather than the content being learned.

Note Taking Behaviours

The high number of participants who responded that they take notes while watching videos infers that beyond basic functionality the specific features of the video player may not be significant in developing active viewing behaviours. The HSC Study Lab video player contains the same basic features found on most online video players; stop, play, pause, search/skip, speed, and closed captions. It did not hold the extra interactive features of the VidDeX player used in the Dodson et al., (2018) study. However, students indicated they participated in the same behaviours e.g., note taking. In fact, note taking, and other viewing behaviours that enabled note taking such as ‘pausing’ or ‘re-watching’ was one of the most significant findings revealed by the questionnaire. Therefore, although Dodson et al., (2018) may be correct when they state “students make use of video player affordances for active viewing” (Dodson et al., 2018, p. 3), it needs to be investigated whether having the tool built into the player increases the opportunity and subsequent engagement. It is possible that having the tool incorporated into the video player simply re-directs an existing student behaviour to a prescribed tool. McGowan et al., (2016) also found students liked to take notes while watching videos and that often this behaviour was linked to an upturn (a peak as this study defined it) in engagement graphs. According to Bonwell and Eison (1991), research has revealed that allowing students to pause during a lecture to take notes significantly increases learning. This finding has been extended to video based learning with a number of studies identifying ‘annotation’ as a form of active learning associated with improved learning outcomes, so it is an important behaviour in online and video-based learning (Chi & Wylie, 2014; Dodson et al., 2018; Ou et al., 2016).

Dropout Rates and Returning to Videos

Within the questionnaire questions 5, 6 and 7 related to exiting or dropping out of a viewing session before a video had finished, which is also referred to as the drop-rate. The learning motivation behind this behaviour is contentious with multiple researchers (Kim et al., 2014; McGowan et al., 2016) having identified it as being caused by low engagement on the part of the student. As a result of their findings, Kim et al., (2014) recommended limiting the length of videos. However, this study revealed other reasons for students dropping out of a video, such as that they have simply found the information they’re looking for. When the responses to question 6 ‘Under what circumstances do you ‘stop’ the video?’ had been coded (Table 8) against the ICAP framework, most students indicated reasons for ending the video that would fall into the category of Active engagement. This infers that exiting a video before it has finished may not be due to a lack of engagement. Kim et al.’s (2014) study found that there was a correlation between video length and dropout rates revealing that students “might feel bored due to (a) shorter attention span or experience more interruption” (Kim et al., 2014, p. 3). However, it appears likely that there are alternative motivations behind this behaviour. The Kim et al., (2014) and McGowan et al., (2016) studies also found that the dropout rate significantly increases when a student is re-watching a video, a finding they state was caused by the student having more specific information needs. The McGowan et al., (2016) and Kim et al., (2014) studies further revealed that when students return to a video a second time, they’re much more targeted in their viewing behaviours, with more re-watching of sections and higher dropouts. The findings of this study reveal this type of behaviour may infer information seeking and/or reviewing of video content (active learning behaviours), rather than a lack of engagement.

The results of this study support those of Lagerstrom et al., (2015), that students often watch a video over multiple sessions rather than in one sitting. The video hosting program provided data visualisations of each viewing session rather than individual viewers, so it was highly likely that some of the sessions were the same student returning and re-watching/completing the video. The pairing of these separate viewing sessions by their unique IP addresses (Figure 8, 9, 10, & 11) reveals behaviours that align with those identified in the Lagerstrom et al., (2015) study where students often returned to re-watch a video. This finding caused Lagerstrom et al., (2015) to question alternative interpretations of dropout rates and video length, including the validity of the 6-minute rule in video length, as recommended by Kim et al., (2014). Lagerstrom et al., (2015) found that returning to complete a video was a common behaviour, and that the aggregate of a student's viewing sessions provide a better understanding of student engagement. The Lagerstrom et al., (2015) study found that dropout rates appeared significant when sessions were viewed individually, but when stitched together, that 90% of students watched almost the whole video. By investigating average dropout rates rather than combining the multiple viewings of individual users, the McGowan et al., (2016) and Kim et al., (2014) studies may have missed this behaviour. This behaviour also speaks to the finding by Lodge et al., (2017) that video-based learning claims like the '6-minute rule', were largely untested.

Due to limitations in the data set harvested from the video hosting service, and therefore available to this study, it was not possible to compare the 870 total individual viewing sessions against one another to identify patterns across individual viewing behaviours. However, question 7 of the questionnaire asked participants if they return to a video (provided they had not finished it the first time) and 74% of respondents self-reported that they do or sometimes do. This provides significant support for Lagerstrom et al.'s (2015) work, and to the reliability of interpretations relating to behaviours identified in the paired (individual) sessions. Both Kim et al., (2014) and Largerstrom et al., (2015) relied solely on trace data for analysis, and as such, student motivations behind observed behaviours were inferred through the data. McGowan et al.'s, (2016) study, by comparison, supplemented their LA method with a questionnaire. McGowan et al., (2016) found in their analysis that there was a significant drop off at the end of a video, and that these areas tended to correspond to a summary or discussion of the next topic. This was interpreted by McGowan et al., (2016) as going against the commonly-held teaching practice of summarising key points at the end of a lecture. However, when asked about this behaviour in the questionnaire, the students stated that while they felt summaries in physical lectures were ‘really necessary’ it was not necessary in videos, as they could rewind when they wanted to review (McGowan et al., 2016). This again reveals a limitation with LA as a sole data method, as it would not have identified the motivation behind the behaviour. Further, however, it infers that dropping out of a video before it has completed is not necessarily a sign of dis-engagement.

An additional finding of this study revealed that when a student abandons a video (irrespective of whether they return) this is still not necessarily an indication of a lack of engagement. When asked ‘Under what circumstances do you ‘stop’ the video?’ (Table 8) the majority of respondents (65%) reported ‘I found what I needed’ and only 16.7% selected ‘It is boring’. This is evidence of an active learning behaviour and similar to that described by Winne (2018) as being a component of self-regulated learning. Winne (2018), states that a student will move on from a chapter or piece of text once they feel they have understood it, in order to limit required effort. This behaviour may be missed if studies rely solely on an LA method that only infers underlying motivations. The examination of individual viewing sessions in this study supported by the results of the questionnaire, infers that students dropping out of viewings sessions may be active learning strategies.

A Review of the ICAP Framework

Studies that did not adopt the ICAP framework as a model for active learning often interpreted behaviours quite differently to those that did. Both McGowan et al., (2016) and Kim et al., (2014) concluded that ‘skipping ahead’ inferred disengagement. Furthermore McGowan et al., (2016) viewed a smooth graph (visualisation of the data) as inferring engagement. The ICAP framework, on the other hand, argues that this behaviour is indicative of active learning (Chi & Wylie, 2014). As discussed in the literature review Lagerstrom et al., (2015) stated that when limited to an LA method motivations behind behaviours could only be inferred, and that more work was needed to understand student intentions. The Dodson et al., (2018) study also revealed examples of skipping ahead, however, by supplementing the data observations with a questionnaire, they were able to identify student motivation behind the behaviour. Students reported that they would often look for information, specifically slides etc. within the video, and then use some kind of note taking tool to take down the information (Dodson et al., 2018). This finding supports an ICAP interpretation of behaviours associated with skipping ahead. The HSC Study Lab investigation also found significant agreement between the results of its questionnaire and the framework of behaviours and underlying motivations as modelled by the ICAP framework.

The ICAP framework focusses only on the product (motoric behaviours) of the student and does not attempt to identify underlying motivations (Stepanek, 2017). This makes it particularly well suited to LA methods that rely on overt actions that generate trace data. However, it also requires that the ICAP framework (and by extension studies that adopt it) rely on additional data to confirm the link between the motoric behaviour and the underlying motivation (Stepanek, 2017). It is possible that there were active learning behaviours invisible to the LA method and furthermore, that there was a disconnect (misalignment) between the actions of the student (those recorded) and the student's underlying intentions. In order to illuminate these limitations a secondary method of a questionnaire was utilised, which revealed that there was significant alignment between observed behaviour and underlying student motivation (Table 2). By supplementing the LA method with a questionnaire this research project has further developed an understanding of student intentions and found that there is a high degree of alignment between the study's findings and the assumptions of the ICAP framework. Therefore, this study was able to contribute to the body of research in support of the ICAP framework as a conceptualisation of active learning. An investigation of the central hypothesis of the ICAP framework that there is a progressive hierarchy of learning outcomes mapped to the modes of engagement (I>C>A>P), was beyond the scope of this study to investigate.

Limitations and Future Work

The study was limited by the data sets available through the Sprout video hosting service. As such it was unable to compare behaviours of large numbers of individual viewers, therefore limiting reliability. This in turn meant it was unable to acquire data on the total number and/or percentage of students who actively interacted with the video as compared to passively. It was also unable to distinguish between behaviours that produced similar patterns in the aggregate data, for example, dropping out and skipping forward. If the data could identify individuals, then it would have been possible to see users who were skipping ahead and re-joining a video viewing session later on – differentiating them from users who were dropping out completely. This also meant it was not possible to determine things like average dropout rate and percentage of video viewed before dropout. As could be seen from the two example users (identified by unique IP addresses) it is likely that students watch videos across multiple sessions, a conclusion also identified by Lagerstrom et al., (2015). As discussed, data analysis revealed that LA could identify patterns of viewing associated with active learning, however, due to limitations in the data set, the analysis was restricted to peaks in viewership caused by students re-watching sections of a video. Additional behaviours could not be identified through the aggregated data such as searching, pausing and individual drop-outs. In order to acquire a complete data set APIs were required to be added to the program which was beyond the immediate resources of this study, however, it is possible that they can be applied for future research.

Due to limitations in the availability of the source data, it was necessary to manually harvest viewership percentages second by second across the video, which made comparison across multiple videos prohibitively time consuming. The ability to compare the data from this video with multiple other videos within the learning environment would strengthen reliability of findings, and the process will be automated for future larger scale studies. However, the video itself had a total number of unique viewings of 870 (at time of data collection), which is a large sample size, and comparison with the questionnaire added reliability to the findings as well as providing additional information around student motivations. The LA method and framework were successful at identifying both Passive and Active engagement (in the form of re-watching sections of the video) but higher-level forms of engagement, those associated with the Constructive mode, were not able to be identified and the questionnaire was relied on to disclose them. According to Naroozi et al., (2019) it is only through the gathering of multimodal data from different sources that identification of complex and invisible metacognitive learning process is possible.

Online learning and video-based learning is now ubiquitous in education, and learning analytics is a growing field of study. Considering the limitations in the data set it is important to extend the study and add to the validity of the findings by applying the methods to richer data sets - either within HSC Study Lab, which will require modification to how the data is collected and made available, and/or to repeat the study in new online environments. The findings of this study were encouraging enough that further research in this area is warranted. Furthermore, considering how new LA is as a method and how little research there has been in online learning environments and video-based learning, there are significant reasons to argue such research would be of benefit to the teaching and learning community. The ultimate goal of future work is the improvement of student learning outcomes, specifically within online learning environments. As Zilvinskis et al., (2017) and Viberg et al., (2018) point out, this has not been the general result of learning analytics interventions.

Conclusion

Online learning environments with no direct teacher instruction or student support are increasing within education (Bowers & Kumar, 2015) and within such environments it is often difficult to assess how learners are making use of the tools available to them (Aleven et al., 2003). Traditional research methods for evaluating such courses can be invasive and time consuming (Winne, 2017) and with that in mind, research into the efficacy of learning analytics for identifying active learning may provide a valuable tool for evaluating online learning environments. If that method were matched with a clear framework for defining meaningful engagement it would be of significant value to course designers and administrators. The research questions this study investigated were to what extent learners participate in active learning in HSC Study Lab, and to what extent LA is an effective method for identifying that active learning? The research project used data harvested from a digital video-player and applied a learning analytics method, mapped to a framework for active learning, in order to identify patterns of viewing associated with active learning.

The results of the analysis support the contention that LA can identify patterns of behaviour associated with active learning in HSC Study Lab. Further, the results of the questionnaire supported the findings of the LA method as students self-reported that they did participate in the behaviours identified, and that they did so for reasons consistent with active learning motivations. It was clear from the results that learners participate in a range of active learning behaviours in online learning and perhaps to a greater extent than previously recognised. As a tool for identifying active learning in video-based learning, this study finds that LA has a role to play, but it is incomplete. It was clear that it could identify re-watching and that this behaviour was associated with the Active engagement mode of active learning. However, what was also clear was that without the questionnaire a lot of the additional active learning behaviours could not have been identified. Student note taking, for example, would fit a mode of engagement higher than simply re-watching, but was not visible in the data. Although further research is required to investigate the outcomes of this study the results so far provide evidence in support of LA having potential for identifying behaviours associated with active learning in online learning environments, and as such, LA could prove to be a valuable tool in the digital environment. Furthermore, the ICAP framework adopted for defining active learning was supported by the results of the questionnaire, and as such there is increased confidence in using it as a conceptual framework for defining modes of engagement associated with active learning.

References

Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. (2003). Help Seeking and Help Design in Interactive Learning Environments. *Review of Educational Research, 73*(3), 277-320. doi:10.3102/00346543073003277

Ali, L., Hatala, M., Gasevic, D., & Jovanovic, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers & Education, 58*.

Allen, I. E., & Seaman, J. (2015). *Grade Level: Tracking Online Education in the United States*: ERIC.

Azevedo, R., Guthrie, J. T., & Seibert, D. (2004). The role of self-regulated learning in fostering students' conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research, 30*(1-2), 87-111.

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In *Learning analytics* (pp. 61-75): Springer.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual review of psychology, 52*(1), 1-26.

Behrens, J. T., & DiCerbo, K. E. (2014). Harnessing the currents of the digital ocean. In *Learning analytics* (pp. 39-60): Springer.

Bonwell, C. C., & Eison, J. A. (1991). *Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports*: ERIC.

Bowers, J., & Kumar, P. (2015). Students' perceptions of teaching and social presence: A comparative analysis of face-to-face and online learning environments. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT), 10*(1), 27-44.

Bowers, J., Kumar, P. J. I. J. o. W.-B. L., & Technologies, T. (2015). Students' perceptions of teaching and social presence: A comparative analysis of face-to-face and online learning environments. *10*(1), 27-44.

Cashion, J., & Palmieri, P. (2002). *The secret is the teacher: The learner's view of online learning*: National Centre for Vocational Education Research.

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in cognitive science, 1*(1), 73-105.

Chi, M. T., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., . . . Stump, G. S. J. C. s. (2018). Translating the ICAP theory of cognitive engagement into practice. *42*(6), 1777-1832.

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist, 49*(4), 219-243.

Chorianopoulos, K., Giannakos, M. N., Chrisochoides, N., & Reed, S. (2014). *Open service for video learning analytics.* Paper presented at the 2014 IEEE 14th International Conference on Advanced Learning Technologies.

Coffrin, C., Corrin, L., de Barba, P., & Kennedy, G. (2014). *Visualizing patterns of student engagement and performance in MOOCs.* Paper presented at the Proceedings of the fourth international conference on learning analytics and knowledge.

Cohen, L. (2007). *Research methods in education / Louis Cohen, Lawrence Manion and Keith Morrison* (6th edition. ed.). London ; New York: London ; Routledge.

de Freitas, S. I., Morgan, J., & Gibson, D. (2015). Will MOOCs transform learning and teaching in higher education? Engagement and course retention in online learning provision. *British Journal of Educational Technology, 46*(3), 455-471.

Dodson, S., Roll, I., Fong, M., Yoon, D., Harandi, N. M., & Fels, S. (2018). *An active viewing framework for video-based learning.* Paper presented at the Proceedings of the Fifth Annual ACM Conference on Learning at Scale.

Fedynich, L., Bradley, K. S., & Bradley, J. (2015). Graduate Students' Perceptions of Online Learning. *Research in higher education journal, 27*.

Giannakos, M. N., Chorianopoulos, K., & Chrisochoides, N. (2015). Making sense of video analytics: Lessons learned from clickstream interactions, attitudes, and learning outcome in a video-assisted course. *International Review of Research in Open and Distributed Learning, 16*(1), 260-283.

Giannakos, M. N. J. B. J. o. E. T. (2013). Exploring the video-based learning research: A review of the literature. *44*(6), E191-E195.

Harris, P., & Walling, D. R. (2014). Policies governing educational technology practice and research. In *Handbook of research on educational communications and technology* (pp. 627-640): Springer.

Jivet, I., Scheffel, M., Specht, M., & Drachsler, H. (2018). *License to evaluate: preparing learning analytics dashboards for educational practice.* Paper presented at the Proceedings of the 8th International Conference on Learning Analytics and Knowledge.

Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). *Understanding in-video dropouts and interaction peaks inonline lecture videos.* Paper presented at the Proceedings of the first ACM conference on Learning@ scale conference.

Kizilcec, R. F., & Halawa, S. (2015). *Attrition and achievement gaps in online learning.* Paper presented at the Proceedings of the Second (2015) ACM Conference on Learning@ Scale.

Kleftodimos, A., & Evangelidis, G. J. S. L. E. (2016). Using open source technologies and open internet resources for building an interactive video based learning environment that supports learning analytics. *3*(1), 9.

Lagerstrom, L., Johanes, P., & Ponsukcharoen, M. U. (2015). *The myth of the six minute rule: student engagement with online videos.* Paper presented at the Proceedings of the American Society for Engineering Education.

Lodge, J. M., Horvath, J. C., Horton, A., Kennedy, G., Venema, S., & Dawson, S. (2017). Designing videos for learning: Separating the good from the bad and the ugly.

Ma, J., Han, X., Yang, J., & Cheng, J. (2015). Examining the necessary condition for engagement in an online learning environment based on learning analytics approach: The role of the instructor. *The Internet and Higher Education, 24*, 26-34.

Marzouk, Z., Rakovic, M., & Winne, P. H. (2016). *Generating Learning Analytics to Improve Learners' Metacognitive Skills Using nStudy Trace Data and the ICAP Framework.* Paper presented at the LAL@ LAK.

Mayer, R. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning, 33*(5), 403-423.

Mayer, R. E. (2004). Should There Be a Three-Strikes Rule against Pure Discovery Learning? *American Psychologist, 59*(1), 14-19. doi:10.1037/0003-066X.59.1.14

Mayer, R. E. (2017). Using multimedia for e-learning. *Journal of Computer Assisted Learning, 33*(5), 403-423.

Mayer, R. E., & Johnson, C. I. J. J. o. E. P. (2008). Revising the redundancy principle in multimedia learning. *100*(2), 380.

McGowan, A., Hanna, P., & Anderson, N. (2016). *Teaching programming: understanding lecture capture YouTube analytics.* Paper presented at the Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education.

Noroozi, O., Alikhani, I., Järvelä, S., Kirschner, P. A., Juuso, I., & Seppänen, T. (2019). Multimodal data to design visual learning analytics for understanding regulation of learning. *Computers in human behavior, 100*, 298-304.

Ou, C., Goel, A. K., Joyner, D. A., & Haynes, D. F. (2016). *Designing Videos with Pedagogical Strategies: Online Students' Perceptions of Their Effectiveness.* Paper presented at the Proceedings of the Third (2016) ACM Conference on Learning@ Scale.

Pardo, A. (2014). Designing learning analytics experiences. In *Learning analytics* (pp. 15-38): Springer.

Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics, 2*(1), 7-12.

Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., & Kloos, C. D. (2015). ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior, 47*, 139-148.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences*. Cambridge: Cambridge University Press.

Scardamalia, M., & Bereiter, C. J. C. o. t. A. (1993). Technologies for knowledge-building discourse. *36*(5), 37-41.

Scheiter, K. (2014). The learner control principle in multimedia learning.

Siemens, G., & Long, P. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review, 46*(5), 30.

Sims, R., Dobbs, G., & Hand, T. (2002). Enhancing quality in online learning: Scaffolding planning and design through proactive evaluation. *Distance education, 23*(2), 135-148.

Spector, J. M. (2001). An overview of progress and problems in educational technology. *Interactive educational multimedia: IEM*(3), 27-37.

Stepanek, N. R. (2017). *Towards Student Engagement Analytics: Applying Machine Learning to Student Posts in Online Lecture Videos.* University of Nebraska at Omaha,

Stepanek, N. R., & Dorn, B. (2017). *Automated Analysis of Lecture Video Engagement Using Student Posts.* Paper presented at the International Conference on Artificial Intelligence in Education.

van Hout-Wolters, B., Simons, R.-J., & Volet, S. (2000). Active learning: Self-directed learning and independent work. In *New learning* (pp. 21-36): Springer.

Verbert, K., Manouselis, N., Drachsler, H., & Duval, E. (2012). Dataset-Driven Research to Support Learning and Knowledge Analytics. *Educational Technology & Society, 15*(3), 133-148. Retrieved from http://simsrad.net.ocs.mq.edu.au/login?url=http://search.proquest.com/docview/1314329982?accountid=12219

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in human behavior*.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*: Harvard university press.

Warren, S. J., Lee, J., & Najmi, A. (2014). The impact of technology and theory on instructional design since 2000. In *Handbook of research on educational communications and technology* (pp. 89-99): Springer.

Winn, W. (2002). Research into Practice: Current Trends in Educational Technology Research: The Study of Learning Environments. *Educational Psychology Review, 14*(3), 331-351. doi:10.1023/a:1016068530070

Winne, P. H. (2017). Leveraging big data to help each learner and accelerate learning science. *Teachers College Record, 119*(3).

Winne, P. H. (2018). Theorizing and researching levels of processing in self-regulated learning. *British Journal of Educational Psychology, 88*(1), 9-20.

Zilvinskis, J., Willis III, J., & Borden, V. M. (2017). An overview of learning analytics. *New Directions for Higher Education, 2017*(179), 9-17.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational psychologist, 25*(1), 3-17.

Ethics Letter of this thesis have been removed as they may contain sensitive/confidential content

Questionnaire

Default Question Block

Thank you for participating in this questionnaire on viewing habits in HSC Study Lab.

Participation is anonymous, and you are free to withdraw at any time. If you choose to, you may provide your email address and go in to the draw for one of ten $50 iTunes vouchers. If you agree to provide your email address, it will not be associated with your responses, will not be used for any purpose other than identifying and notifying prize winners and will be destroyed once prize winners are selected and notified.

Any information or personal details gathered in the course of the study are confidential, except as required by law. No individual will be identified in any publication of the results. A summary of the results of the data can be made available to you on request – please email Eamon Vale if you wish to receive a summary of the data.

Participation in this study is entirely voluntary: you are not obliged to participate and if you decide to participate, you are free to withdraw at any time without having to give a reason and without consequence.

Where do you watch the video-based lessons?

Home

School

Other. Please detail

Do you watch the video-based lessons on your own or with others?

On my own

With others

If 'yes' who do you watch the video-based lessons with?

Classmate

Teacher

Other. Please detail

When you start a video how often do you finish it?

Always

Usually

Sometimes

Never

Under what circumstances do you 'stop' the video?

It is boring

It is too difficult

I found what I needed

Other. Please detail

Do you return to finish the video later (yes, no)?

Yes

No

Sometimes

Do you take notes while watching the video (yes, no, sometimes)?

Yes

No

Sometimes

Where do you take notes?

Computer

Notebook

Other. Please detail

Do you pause the video to take notes (yes, no)?

Yes

No

Sometimes

Do you discuss the content of the video lessons with another person (yes, no)?

Yes

No

Sometimes

Who do you discuss the videos with?

Classmate

Teacher

Other. Please detail

While watching a video do you skip back to 're-watch' parts of it?

Always

Sometimes

Never

Why do you re-watch the videos?

It was confusing

It was interesting

Other. Please detail

Please indicate what video features you use when watching the videos?

Pause button

Scan forward/backward

Speed

Closed captions

Please comment on the strengths and weaknesses of video-based lessons.

When watching videos which ‘strategy/s’ applies best to you (more than one response allowed)?

Watch everything at once

Take notes while watching

Watch the video with a classmate

Search the video for 'important points'

Other. Please detail

Please indicate if there are additional features or tools you would like in or related to the videos to help you learn better.

Thank you for participating in the HSC Study Lab video-based learning questionnaire.

If you wish to go into the draw for a $50 iTunes voucher, please enter your email address in the text field.

By entering your email address you are waiving your anonymity, however, your email address will not be associated with the responses from your questionnaire, will not be used for any other purpose, or provided to any other parties, it will not be stored with the questionnaire data, and will be destroyed once prize winners are selected and notified.