# At the heart of a diverse technology: Applying a realist evaluation methodology to a university live streaming programme

**Mike Bryant**
Monash University
Australia

**Colin Simpson**
Monash University
Australia

**Alexander Whitelock Wainwright**
Monash University
Australia

**Trevor Wood**
Monash University
Australia

**Kris Ryan**
Monash University
Australia

Researching web-based lecture technologies in higher education is a complex undertaking (Morris, Swinnerton, & Coop, 2019), making experimental trials and summative evaluations impractical. The implications of existing literature in this area are therefore uncertain (Kay, 2012; O'Callaghan, Neumann, Jones, & Creed, 2017). In this paper we report on the use of an alternative research design, realistic evaluation (King, Dawson, Rothberg, & Batmaz, 2017; Pawson and Tilley, 1997), to gain insight into the impact of a university lecture live-streaming initiative. The initiative provided synchronous and asynchronous access to video of 129 weekly lectures using two different platforms. Surveys of 306 students and 49 staff, a focus group of five students, and an exploratory study of trace data of 359 students indicated that student access to the videos was lower and more idiosyncratic than expected. The ability to identify useful content, appropriate instructional methods, and participation by peers were reported to encourage students to become involved. Our results confirm the importance of specific contexts and forms of behaviour in encouraging beneficial lecture video use. We also show time of initial view to be a promising way to study students' strategies for using lecture video content.

Keywords: realistic evaluation, learning analytics, web-based lecture technologies

## Introduction

Realist evaluation, as described by Pawson and Tilley (1997), adopts the critical realist paradigm (Carlsson, 2012) to propose a theory-based model for evaluating social initiatives or programmes. 'Theory' is defined by Pawson and Tilley as a set of beliefs about the conditions and processes/mechanisms likely to produce a regular observable effect; they argue that theories underlie the design of all social initiatives and can be tested and refined through evaluation. Rather than aiming to establish generalisable *post-hoc* causal explanations, realist research designs in evaluation studies seek to engage and inform policymakers by identifying *how*, *why* and *for whom* initiatives bring about change. Realist evaluation as an approach was quickly embraced by researchers working in fields such as health and medical education (Ogrinc & Bataldan, 2009; Marchal, van Belle, van Olmen, Hoeree, & Kegels, 2012; Hewitt, Sims, & Harris, 2013). More recently, the use of realist evaluation has also been advocated for educational technology research (Sorinola, Thistlethwaite, Davies, & Peile, 2015; Stohr & Adawi, 2018; McFaul & Fitzgerald, 2019) because of its potential to address the practical complexities of educational ecosystems (Ellis & Goodyear, 2019). Using this approach, our goal here is to show that effective, impactful evaluation of complex activity is not only achievable but well within reach for educational technology programme teams, as long as specific behaviours and specific contexts are placed at the heart of the enterprise.

In our study, a realist evaluation approach was adopted to establish how and in what circumstances a lecture live-streaming initiative at a large, research-intensive Australian higher education institution, Monash University, brought about beneficial changes. The research team included staff responsible for designing and implementing lecture live-streaming; only the first phase of evaluation is reported. The study was motivated by a desire to explore the efficacy of realist evaluation in the context of a complex educational technology initiative. Realist evaluation uses diverse methods to explore the specific social conditions and forms of behaviour that are associated with the success of a particular initiative; it seeks to improve specificity of understanding and ability to predict outcomes across multiple studies. Accordingly, both in conducting the research and for the purposes of reporting our findings in this paper, the research team followed the cycle of activity described in Pawson and Tilley (1997, chapter 4): investigate the initial design or "specification" of the initiative; generate theories and hypotheses; design and conduct observations; refine theories and discuss implications.

The rationale for our chosen approach has two key elements. Firstly, there is a need to improve the evaluations that are currently undertaken of educational technology initiatives in higher education (Nordmann, Calder, Bishop, Irwin, & Comber, 2018). The need is discussed at length in Selwyn (2014) and King, Dawson, Rothberg, & Batmaz (2016). Whilst there is already a substantial body of literature relating to web-based lecture technologies (WBLT), its conclusions are mixed and their implications are uncertain (Kay, 2012; O'Callaghan, Neumann, Jones, & Creed, 2017). We do not necessarily agree with the conclusion of Nordmann and colleagues that this is a situation best remedied through further empirical work, coupled with rigorous meta-evaluation or further research synthesis. Instead, we suggest that members of the research team and our equivalents in other institutes are faced with a classic evaluator's dilemma: a plethora of evidence and apparently conflicting findings on one hand, and a pressing need to measure the value of initiatives on the other (Pawson & Tilley, 1997). Data collection about WBLT is a highly complex undertaking (Morris, Swinnerton, & Coop, 2019) and institutional and practical constraints restrict the feasibility of experimental trials and full-scale summative and/or independent evaluations (O'Callaghan et al., 2017). In this context, we consider that more research by itself may not help; it could even intensify the difficulty. What is most likely to be required in our view is a re-think of educational technology evaluation and its relationship to research. Our study is intended as a contribution towards that goal.

Secondly, there is an often-stated view that theory has not kept pace with the use of educational technology (King, Dawson, Rothberg, & Batmaz, 2017; Morris et al., 2019). One possible reason for this is that educational change involves interdependent changes in behaviour (student and staff), in curricula and in technical systems, motivated by the beliefs and values that prevail in a particular educational community (Fullan, 2016). Fullan's work suggests that theories of technology are unlikely by themselves to offer a worthwhile explanation as to what is happening in universities.

We therefore concur with Whitworth's (2012) call for frameworks of evaluation that permit attention to the political and institutional context of educational technology initiatives. The approach of realist evaluation provides one such framework, offering a promising alternative model for integrating research within the development and evaluation of technology initiatives that is grounded in critical realism.

We begin by considering the findings of King et al. (2017), as this research team also adopted a realist evaluation approach in a recent and comparable study of a lecture video initiative at a research-intensive university in the United Kingdom. Modelling their work closely on Pawson and Tilley (1997), King et al. concentrated on explaining observed regularities in the use of lecture video via theory development. The authors use the term "theory" in the same way as Pawson and Tilley to mean "propositions about how mechanisms are fired in contexts to produce outcomes" (1997, Chapter 4) refined in a cyclical way via data collection and analysis over multiple studies. In other words, they attempted to develop ideas about the use of lecture video and the technical performance of the video platform, so as to reliably explain – and therefore ultimately predict – video access patterns, as well as informing future research.

We summarise the theories resulting from the work of King et al. (2017) as follows:

i. Students are encouraged to utilise video capture of scheduled teaching when this opportunity appears aligned to their study objectives and personal preferences for learning, and also when the video is accessible and of acceptable quality.
ii. In departments where the use of video capture of scheduled teaching is embedded, students become familiar with its use, although this familiarity can lead to disillusionment or avoidance if students do not acquire the skills to use video effectively or regard its use as an unsuitable vehicle for learning.
iii. In units where video capture of scheduled teaching has been relatively successful, staff report that they are encouraged to use the technology by student demand and for equity reasons.

We accept, though not entirely without reservation, the use of lecture video access patterns by King et al. (2017) as a proxy for the "outcomes" of a lecture video initiative. However, Pawson and Tilley (1997) stipulate that an empirically observed pattern will be affected by *specific forms of behaviour* in *specific contexts*. We therefore sought to develop a realist research design for lecture video that considers when and how, as opposed to simply how often, video is accessed. Critically, the initiative that is the basis of research reported in this paper provided students with the opportunity for synchronous (live-stream) access to lecture video alongside the more familiar offering of recorded lectures. Consequently, the full range of possible access times could be investigated. We believe ours is the first use of a context-sensitive realist design in the study of web-based lecture technologies (WBLT). Our research therefore constitutes a distinctive contribution to the emerging literature on realist approaches to the study of higher education technology initiatives.

## Initial investigations: live-streaming at Monash University

Up to and including 2017, video capture of scheduled teaching activities was made available to students in some Monash University units at the request of the academic or the department. From time to time, individual disciplines also made class video capture footage available synchronously in order to experiment with the potential offered by live-streaming.

A major initiative known as Monash LIVE commenced at the start of 2018. This involved the creation of a scheduling option for students to attend lectures via live-streaming in units with a relatively high number of students enrolled. Ninety-one units (76 undergraduate and 15 postgraduate units) provided a scheduled live-streaming option to students as part of the LIVE initiative during the first major teaching period of that year. 129 average weekly scheduled teaching activities were live-streamed, representing 3.66% of the total number of lectures on Monash's Australian campuses. Two different video platforms were made available, described here as platform A and platform B; 14 lecture rooms were upgraded with new audio-visual equipment to support lecture live-streaming. At the same time, a separate project was undertaken to clarify and update the nomenclature, timetable codes and definitions used to describe scheduled teaching activities.

The research team searched in papers from the initiative steering committee, scheduling data (in 2017, 2018 and 2019) and information provided to students and staff, comparing the specification of Monash LIVE with that of the initiative described in King et al. (2017). Both initiatives used policy change and student demand to introduce WBLT across the institution, formalising and standardising existing localised practices. While the other initiative expected that students would learn more effectively from lectures when recordings were made available for review and revision, Monash LIVE anticipated that pressure on lecture venues would be reduced when the possibility was created for students to attend and participate in lectures synchronously online. Both initiatives expected widespread take-up by students; in the case of Monash LIVE there was an additional expectation that students and staff would quickly become familiar with interacting using text questions and comments during the time period scheduled for the lecture.

Further to the theories about lecture video use identified by King et al. (2017), we identified two additional initial theories specifically associated with Monash LIVE. These were:

iv. Students and staff quickly become accustomed to large-cohort interactive synchronous whole-class teaching that spans in-person and remote attendance.
v. The presence of a member of staff with the role of 'moderator', working alongside the lecturer to monitor and respond to text questions and comments from live-streaming students, mitigates the practical challenges of online interaction.

We proceeded to develop a research design that would allow us to refine both sets of theories and to better understand the explanatory potential of context-sensitive data within the overall framework of a realist approach.

## Designing and conducting observations of Monash LIVE

Our research relates only to the first phase of the Monash LIVE initiative in the first main teaching period of 2018. Because there is little consensus and considerable fragmentation in existing WBLT literature (see above), we chose not to develop lines of enquiry for the study through a literature review. Some theory has been established, but not, so far, informed by observations in specific contexts. In a theory-poor environment, Pawson and Tilley's matrix of realist designs (1997, chapter 4, table 4.1) recommends the use of exploratory strategies, typically those involving concurrent quantitative and qualitative data collection. For this reason, we used multiple methods to gather and compare data about the specific contexts and uses of WBLT in the Monash LIVE initiative. We surveyed 306 students and 49 staff, held a focus group ($n$ = 5), and carried out a pilot study of trace data of video usage by 359 students in one first year unit. Next, we compared the results with those in the existing literature in order to refine theories i. - v. Our study concluded with a review of major findings and their implications, reflection on the efficacy of the approach taken, and consideration of next steps.

## Two research questions informed these activities:

*a) What insights can a context-sensitive realist evaluation methodology provide, through improved theory, that might support better programme design and/or institutional practice in a WBLT initiative based on live-streaming?*

*b) What can the findings of a) suggest about the benefits and limitations of realist evaluation as a means of assessing the impact of educational technology initiatives?*
We invited all students in the 91 LIVE units, including those not enrolled in the live-streaming lecture option, to complete a survey. Student survey items were adapted from those reported in Richardson, Dunn, McDonald, and Oprescu (2014). Three underlying constructs were identified: overall opinion of the initiative, benefits to learning, and technical usability. The development of the student survey is reported in detail in Bryant, Francis, Ryan, Wood, and Zhang (2019).

A total of 596 students responded, of whom 306 provided complete responses. The scale was a seven-point Likert scale where 1 = "strongly disagree" and 7 = "strongly agree". Respondents typically held a positive overall view of live-streaming ($M$ = 5.48, $SD$ = 1.44 for the "overall opinion" construct), regarded the benefit to learning as moderate ($M$ = 4.21, $SD$ = 1.45) and reported a positive overall view of usability ($M$ = 5.01, $SD$ = 1.26). Respondents who provided an email address were invited to attend a focus group to further explore the survey questions. A transcript of the focus group and free-text comments from the staff survey were coded by two members of the team independently and an aggregated set of codes produced, with the aim of representing all sentiments related to the three constructs measured in the survey.

There were 50 complete responses to the staff survey, adapted from the survey reported in Buchanan, Sainter, and Saunders (2013). The scale was an five-point Likert scale where 1 = "strongly disagree" and 5 = "strongly agree". One response related to a previous teaching period and was removed. The overall perception of live-streaming was, on average, neutral: $M$ = 3.00, $SD$ = 1.26 for the item "Overall I was satisfied with the experience of using live streaming software". Most respondents utilised the open comment fields. Accordingly, and because of the small sample size, we analysed free-text rather than numerical responses.

Students in the focus group indicated that when they were able to identify video that was likely to be valuable for the purpose of study and view that content in a setting that promoted concentration, they placed a high value on the use of the platforms for learning. For example, a student reported intentionally leaving campus so as to be able to view part of a lecture in a quiet location. We suggest that this sentiment may explain the typically high opinion of the initiative reported by respondents to the student survey. Staff indicated that where live-streaming was valued, this was either on equity grounds (to provide a degree of access to those students unable to attend in person, as suggested by King et al., 2017 in theory c.) or because it could support the goals of a specific unit or stage of the course.

One factor mitigating against the perceived benefit to learning for students was the absence of live video of the lecturer, reducing the ability to pick up non-verbal cues such as gestures or use of the pointer. Where internet, software or hardware problems restricted access, this also affected perceived educational value. No students in the focus group and few students in the survey ($M$ = 2.31 for the relevant item) identified their own skill level or difficulty with using the platform as a barrier. Factors that staff reported as reducing the perceived value of the initiative were scepticism about its aims, insufficient training and support, and insufficient time to prepare. The complexity of using live-streaming as part of large-class lecturing was also reported by some staff as having reduced its effectiveness.

Both student and staff respondents reported that synchronous access and participation had depended heavily for their usefulness on the instructional methods employed and students' responses to them. Students identified instructional choices they saw as tokenistic and low engagement from peers as factors likely to discourage synchronous participation. Staff stated that the level of student engagement in live-streaming, which some believed had increased the likelihood of absence, was low and insufficient to justify the time required to implement the initiative within their unit. Latency (time delay) – built in to the platform and/or the result of buffering – was identified in the staff free-text comments as an inhibitor of synchronous platform use. The use of moderators was reported as inconsistent. A total of 20 of the 49 respondents to the staff survey stated either that no moderator was assigned, or that the moderator did not have an active role because of low synchronous engagement by students. The intention had been that synchronous use of platform A would be the norm, with students and staff moving to platform B only in the event of technical problems. This was not what occurred in practice. Some scheduled teaching activities occurred in venues where only platform B was available. Also, students in the focus group reported idiosyncratic patterns of technology access, with personal schedules, preference and experience of platform performance all playing a role in influencing when content was accessed and how.

Time of access was investigated further in a small-scale exploratory study of a single unit using trace data from platform A. This unit, a first year psychology unit with two weekly lectures (25 videos in total), was chosen because the lecture videos, comprising introductory overviews of a range of topics such as developmental and

biological psychology, were formally and stylistically similar. We analysed students' first views (the time when the first view of each video occurred in relation to the time of the lecture). 996 students enrolled in the unit before semester, with 243 and 95 respectively selecting the live-stream option for the two lectures. 359 students (36.04%) viewed video content. We classified first views as to whether they occurred at the time of the lecture; within one day; within one week; within two weeks; or more than two weeks later. The data were then subject to latent class analysis for the purpose of exploring strategies enacted by students when using live-streaming systems; a three class solution was selected as the candidate model. The composition of the three classes was: 302 students in class 1 (84.12%), 11 students in class 2 (3.06%), and 46 students in class 3 (12.81%).

**Table 1. Descriptive Statistics for the Three Class Solution**

| Class | Live Streaming | | Within 1 Day | | Within 1 Week | | Within Two Weeks | | Greater than Two Weeks | |
|---|---|---|---|---|---|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| 1 | .77 | .90 | .15 | .37 | .58 | .79 | .09 | .29 | .10 | .40 |
| 2 | 6.55 | 2.11 | .36 | .67 | .73 | 1.19 | .18 | .41 | .27 | .65 |
| 3 | .20 | .54 | .89 | .99 | 3.26 | 2.13 | 1.33 | 1.25 | 1.65 | 2.41 |

Table 1 provides the mean and standard deviation values for each of the five variables created from the trace data (time of the lecture; within one day; within one week; within two weeks; or more than two weeks later) across the three identified classes. For example, Class 1 demonstrated limited engagement with the recorded lecture videos based on mean values (e.g., .77 views, on average, for initially live streaming a lecture). Initial viewing behaviour for Class 2, on average, appeared to consist of live-streaming ($M$ = 6.55), whilst all other viewing behaviour occurred less frequently (e.g., $M$ = .18 for within two weeks). Class 3's initial viewing behaviour, based on the presented mean values, generally took place within one week ($M$ = 3.26) or more of the lecture being recorded ($M$ = 1.33 and 1.65 for within two weeks and greater than two weeks, respectively).

This study of trace data supported our emerging understanding that, in contrast to the intention of the Monash LIVE initiative, students' first views of video content were often not synchronous. Instead, we found support for distinct patterns of use. Only a small minority of students, the smallest class (Class 2) followed the guidance provided by Monash LIVE in that their typical first view was under live-streaming conditions. A larger group (Class 3) had first-view behaviour that could be characterised as 'strategic', being one day or more after the time of the lecture on average. We also showed that students made only moderate overall use of the video content available to them in this unit, with mean first views of videos being below 7 in each time category for each of the three latent classes. It should be remembered that our trace data study investigated first views, not duration of view or number of repeat viewings, and only within a single unit and platform.

## Theoretical development and comparison with existing literature

We refined the initial theories as follows:

vi. Students quickly become accustomed to a new model of incorporating video content into their units.
vii. Students are more likely to access video content when it is clearly identifiable as useful for study. In this situation, they will typically seek out an optimal time and physical location to view the content.
viii. Specific instructional methods and the perceived engagement of peers are factors that, in combination, encourage students to participate synchronously.
ix. Perceived high student engagement in live-streaming, equity of access, course and/or unit alignment, and low latency are important factors in encouraging staff to utilise the technology.

Theories vi. - ix. offer several clarifications and improvements to the theory taken from the earlier study of King et al. (2017). The earlier theories i. and ii. proposed that students tend rapidly to develop a superficial acquaintance with WBLT as part of their technical and social environment – as something 'ready-at-hand' – without either fully understanding or accepting their educational purpose in the context of their degree course, sometimes in turn leading to disaffection. Our results indeed indicate low to moderate student use of WBLT based on staff perception

as shown in the free-text survey data, and moderate use in one specific context as indicated in the trace data pilot study. However, we found no evidence that the absence of widespread, wholehearted student access to lecture video at the time of the lecture or subsequently was caused by a general phenomenon of alienation or disillusionment with the technology. Instead, the results both of the student focus group and the staff survey support the team's initial expectation that student motivation to access video is closely associated with specific contexts and forms of behaviour 'in-play' at the time of access. The trace data demonstrates that student first-view behaviours in the selected unit followed several distinct patterns. Theories vi., vii. and viii. reflect these findings, while theory ix. develops the earlier theory iii. about factors influencing staff to use WBLT.

We next investigated whether findings from the extensive broader literature on WBLT could assist with theory development. Several previous studies that utilise system data have found WBLT use by students to be lower than expected. King et al. (2017) found that individual videos were watched by 14% of the student cohort on average across the institution; at unit level, Edwards and Clinton (2019) and Elliott and Neal (2016) both report that a minority of the total number of students enrolled viewed lecture video content when very short views are removed from the data. By contrast, Morris et al. (2019) reported that the proportion of students across the whole institute viewing WBLT at least once reached 81% at the end of a four-year period. Although we did not regard these results as theory-refining in themselves, we noted that our finding of moderate video access by students was a common theme in the literature.

Many previous studies also show that students attribute a high overall value to WBLT, although few specifically consider synchronous access (Chapin, 2018). Self-reported student perceptions of the educational value of WBLT in previous studies strongly support the indication in our results that students see these technologies as providing different affordances to those of in-person lecture attendance. For example, students have reported that they value being able to pause and skip backwards to interrogate complex concepts as well as skipping forward to access specific content strategically (Chapin, 2018; Cilesiz, 2015; Karnad, 2013; Nordmann, 2018; O'Callaghan et al., 2017). Dommett, Gardner, and van Tilburg (2019) found that WBLT access reduced anxiety for the students in their study; a reduction in stress and efficient use of study time is also proposed by Danielson, Preast, Bender, and Hassall (2014) as the most plausible explanation for the positive relationship between number of views and standardized test scores in the lecture-based units they studied. The interviews reported in Cilesiz (2015) indicate that students may progressively become acculturated into the use of WBLT for study. We found these studies to be of significant explanatory value in relation to our theoretical framework; we reflect on their importance in the section below.

We believe that ours is the first study to consider time of first access as a possible indicator of different student behaviour patterns. However, there is other existing evidence that student use of WBLT is variable across the student cohort (Bos, Groeneveld, van Bruggen, & Brand-Gruwel, 2016; Edwards & Clinton, 2019; Leadbeater, Shuttleworth, Couperthwaite, & Nightingale, 2013; Williams, Birch, & Hancock, 2012). Brooks, Erickson, Greer, and Gutwin (2014) found evidence at individual unit level of accesses by different groups of students across different weeks in a teaching period. Danielson et al. (2014) and Nordmann et al. (2018) both found that student access within a department varied by the type of unit, although neither study separates instructional method from unit design. Owston, Lupshenyuk, and Wideman (2011) concluded – based on self-reported use and unit marks – that lower achieving students were more likely to be helped by WBLT, although it seems equally plausible that students who failed to make effective use of the videos may have earned lower marks.

Staff reservations about the introduction of WBLT are reported in Dona, Gregory, and Pechenkina (2017), Freed, Bertram, and McLaughlin (2014), Morris et al. (2019), and Taplin, Kerr, and Brown, (2014). However, very little attention has been paid in existing literature to the goals of staff in using the technologies or staff alignment with higher level institutional strategic aims related to their use. This gap is addressed by the theories we present above.

Few previous studies have explored patterns of viewing behaviour by students in detail using system data – for example, by considering which elements of the video were viewed and what type of action (e.g. rewind or skip forward) led to the view. This remains a gap in the literature and an area for further investigation.

## Discussion and implications

In relation to our research questions:

*a) What insights can a context-sensitive realist evaluation methodology provide, through improved theory, that might support better programme design and/or institutional practice in a WBLT initiative based on live-streaming?*

The evaluation of the first phase of the Monash LIVE initiative allowed us significantly and usefully to refine theories presented by King and colleagues (2017). Although the students we surveyed attributed a high value to live-streaming overall, our focus group found evidence that students do not think about WBLT in the same way as they think about in-person lecture attendance. Instead, value was placed on the ability to identify relevant content and view this in a setting that promoted learning, using the specific affordances of video. Worthwhile synchronous access to and participation in WBLT were associated by students with a combination of appropriate instructional methods and a high level of engagement amongst peers. Our pilot study of trace data supported the findings of the student focus group and confirmed that multiple distinct patterns of first view prevailed in the unit studied, which was a first year introductory science unit. We hypothesise that within a particular initiative, patterns in viewing behaviour are likely to vary according to the design of a unit, the instructional methods adopted and students' existing acculturation into the use of video, and recommend these as areas for future research in the realist paradigm. For staff, student engagement, equity of student access, alignment with course and/or unit aims and low-latency video were reported as encouraging (and their absence inhibiting) efforts to use the technologies synchronously.

Our research design employed a student focus group, qualitative analysis of staff survey responses and study of trace data within a single unit to improve understanding of the specific conditions for students' successful use of video. These methodological choices permitted details of the time of first view, physical location, instructional conditions and video type to be considered as part of theory formation, in a way that would not have been possible had the team relied on aggregated survey data and/or video accesses. Time of initial view was shown to be a promising way to investigate students' strategies for using WBLT. Overall, although there are strong indications in our evaluation that the intention of the initiative for large-scale synchronous engagement was not enacted in practice, our results confirm the importance of specific contexts and forms of behaviour in encouraging beneficial lecture video use.

In the light of these insights, we intend to refine the design of the initiative in various ways. Using the outcomes of a separate project to clarify timetable codes and definitions for different scheduled activity types, the team plans to change the guidance to staff to clarify why, under what circumstances and in which venues live-streaming is made available to students. A series of hardware and software improvements will be undertaken to minimise latency and in-venue technical difficulty. In-venue visual indicators where there is no video stream or the microphone is muted, and the option for academic teams to provide live video of the lecturer, will be provided. Staff will be surveyed at the beginning of the teaching period so that they are able to self-identify a requirement for more training or technical support.

In the next stage of the evaluation, we plan to test and improve theory vii. above by identifying contexts through, and mechanisms by which, students are able to identify video content as potentially valuable for study, and use this information to inform unit design. Our intention is also to investigate contexts in which synchronous engagement is moderate or high and those in which it is less so, exploring specific instructional methods that may encourage students to take part in real-time so as to further refine theories vii. and ix. Finally, we plan to continue to investigate the different strategies adopted by students when using WBLT, expanding our study of trace data across both platforms and refining our survey using items from Danielson et al. (2014).

*b) What can the findings of a) suggest about the benefits and limitations of realist evaluation as a means of assessing the impact of educational technology initiatives?*

We believe the use of realist evaluation has provided us with three key advantages in carrying out the research reported in this paper. Firstly, our chosen design sought to understand the beliefs that had informed the design of this specific initiative, find out to what extent they were justified, and propose refinements or alternatives. There was no intention to establish the general benefits of lecture live-streaming in higher education or to compare live-streaming as an innovation with existing practices of in-person lecture attendance. We were thus able to avoid formidable theoretical and methodological challenges that we expect would be associated with these alternative research approaches. Secondly, this type of design encouraged us to concentrate on specific conditions in which WBLT were used and how these might compare with those of previously reported studies. We were therefore able to add to the research reported in King et al. (2017), for example by exploring a context that involved synchronous use of WBLT. Thirdly, unlike experimental designs, the realist model was *feasible* within the practical constraints governing Monash LIVE. Rarely, in our experience, is there time at the start of such initiatives to develop an approach to evaluation or to carry out a review of literature that might otherwise usefully inform practice. Although we would reserve judgement for future phases of our work, on our current indications we believe this type of research design may have much to offer to research teams working in similar settings, allowing them to get 'to the heart' of diverse technologies within the resources available to them.

One possible limitation on the evidence to date is the extent to which a realist evaluation research design is able to provide a satisfactory account of the experiences of people whose lives are affected by a social initiative. We avoided some of the language associated with realist accounts of the use of technology by individuals, finding it reductive and hard to apply. For example, was the focus group student's choice to leave campus to watch a lecture an example of a "mechanism" or a "context"? Studying "what works" also carries with it a theoretical difficulty of its own. As shown by Biesta (2007), the questions need to be asked: works for whom, and works to do what? Difficulties associated with these questions arose in our research. At times there was a need more clearly to distinguish the benefits of the technology for an individual from the accomplishment of the goals of the programme or the institution, and more fully to investigate different possible ways in which the technology might "work for" or assist learning. Finally, the methodological pluralism of realist designs as described by Pawson and Tilley (2007) carries with it some risk of a lack of rigour. For instance, selection bias may have influenced the results of our surveys, as some staff and students chose not to respond. Had we concentrated exclusively on surveys, we may have found ways to reduce this possibility. We recommend that future research with a similar design takes all three potential limitations into account.

## References


Biesta, G. (2007). Why "what works" won't work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory,* 57(1), 1-22. https://doi.org/10.1111/j.1741-5446.2006.00241.x

Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The revised two factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology, 71*(1), 133–149. https://doi.org/10.1348/000709901158433

Bos, N., Groeneveld, C., van Bruggen, J., & Brand-Gruwel, S. (2016). The use of recorded lectures in education and the impact on lecture attendance and exam performance. *British Journal of Educational Technology, 47*(5), 906–917. https://doi.org/10.1111/bjet.12300

Brooks, C., Erickson, G., Greer, J., & Gutwin, C. (2014). Modelling and quantifying the behaviours of students in lecture capture environments. *Computers & Education,* 75, 282–292. https://doi.org/10.1016/j.compedu.2014.03.002

Bryant, M., Francis, M., Ryan, K., Wood, T., & Zhang, S. (2019). Student voice as feedback: An instrument to measure student perceptions of live-streaming. Presented at the Western Australian Teaching and Learning Forum.

Buchanan, T., Sainter, P., & Saunders, G. (2013). Factors affecting faculty use of learning technologies: Implications for models of technology adoption. *Journal of Computing in Higher Education, 25*(1), 1–11. https://doi.org/10.1007/s12528-013-9066-6

Carlsson, S. A. (2012). The potential of critical realism in IS research. In Y. K. Dwivedi, M. R. Wade, & S. L. Schneberger (Eds.), *Information systems theory: Explaining and predicting our digital society* (Vol. 2, pp. 281–303). New York: Springer. Springer eBooks.

Chapin, L. A. (2018). Australian university students' access to web-based lecture recordings and the relationship with lecture attendance and academic performance. *Australasian Journal of Educational Technology, 34*(5), 1–12. https://doi.org/10.14742/ajet.2989

Cilesiz, S. (2015). Undergraduate students' experiences with recorded lectures: Towards a theory of acculturation. *Higher Education, 69*(3), 471–493. https://doi.org/10.1007/s10734-014-9786-1

Danielson, J., Preast, V., Bender, H., & Hassall, L. (2014). Is the effectiveness of lecture capture related to teaching approach or content type? *Computers & Education,* 72, 121–131. https://doi.org/10.1016/j.compedu.2013.10.016

Dommett, E. J., Gardner, B., & van Tilburg, W. (2019). Staff and student views of lecture capture: A qualitative study. *International Journal of Educational Technology in Higher Education, 16*(1). https://doi.org/10.1186/s41239-019-0153-2

Dona, K. L., Gregory, J., & Pechenkina, E. (2017). Lecture-recording technology in higher education: Exploring lecturer and student views across the disciplines. *Australasian Journal of Educational Technology, 33*(4), 122–133. https://doi.org/10.14742/ajet.3068

Edwards, M. R., & Clinton, M. E. (2019). A study exploring the impact of lecture capture availability and lecture capture usage on student attendance and attainment. *Higher Education, 77*(3), 403–421. https://doi.org/10.1007/s10734-018-0275-9

Elliott, C., & Neal, D. (2016). Evaluating the use of lecture capture using a revealed preference approach. Active *Learning in Higher Education, 17*(2), 153–167. https://doi.org/10.1177/1469787416637463

Ellis, R. A., & Goodyear, P. (2019). *The education ecology of universities: Integrating learning, strategy and the academy*. Routledge.

Freed, P. E., Bertram, J. E., & McLaughlin, D. E. (2014). Using lecture capture: A qualitative study of nursing faculty's experience. *Nurse Education Today, 34*(4), 598–602. https://doi.org/10.1016/j.nedt.2013.06.021

Fullan, M. (2016). *The new meaning of educational change* (5th ed.). New York, NY: Teachers College Press.

Hewitt, G., Sims, S., & Harris, R. (2012). The realist approach to evaluation research: An introduction. *International Journal of Therapy and Rehabilitation, 19*(5), 250–259. https://doi.org/10.12968/ijtr.2012.19.5.250

Karnad, A. (2013). Student use of recorded lectures (p. 24). Retrieved from London School of Economics website: http://eprints.lse.ac.uk/50929/1/Karnad_Student_use_recorded_2013_author.pdf

Kay, R. H. (2012). Exploring the use of video podcasts in education: A comprehensive review of the literature. *Computers in Human Behavior, 28*(3), 820–831. https://doi.org/10.1016/j.chb.2012.01.011

King, M. R. N., Rothberg, S. J., Dawson, R. J., & Batmaz, F. (2016). Bridging the edtech evidence gap: A realist evaluation framework refined for complex technology initiatives. *Journal of Systems and Information Technology, 18*(1), 18-40. https://doi.org/10.1108/jsit-06-2015-0059

King, M. R. N., Dawson, R. J., Rothberg, S. J., & Batmaz, F. (2017). Utilizing a realist evaluative research approach to investigate complex technology implementations. *Journal of Systems and Information Technology, 19*(1/2), 22–41. https://doi.org/10.1108/jsit-04-2017-0027

Leadbeater, W., Shuttleworth, T., Couperthwaite, J., & Nightingale, K. P. (2013). Evaluating the use and impact of lecture recording in undergraduates: Evidence for distinct approaches by different groups of students. *Computers & Education*, 61, 185–192. https://doi.org/10.1016/j.compedu.2012.09.011

Marchal, B., van Belle, S., van Olmen, J., Hoerée, T., & Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation, 18*(2), 192–212. https://doi.org/10.1177/1356389012442444

McFaul, H., & FitzGerald, E. (2019). A realist evaluation of student use of a virtual reality smartphone application in undergraduate legal education. *British Journal of Educational Technology, 0*(0). https://doi.org/10.1111/bjet.12850

Morris, N. P., Swinnerton, B., & Coop, T. (2019). Lecture recordings to support learning: A contested space between students and teachers. *Computers & Education*, 140. https://doi.org/10.1016/j.compedu.2019.103604

Nordmann, E. (2018). The value of capture: Taking an alternative approach to using lecture capture technologies for increased impact on student learning and engagement.

Nordmann, E., Calder, C., Bishop, P., Irwin, A., & Comber, D. (2018). Turn up, tune in, don't drop out: The relationship between lecture attendance, use of lecture recordings, and achievement at different levels of study. *Higher Education*, 1–20. https://doi.org/10.1007/s10734-018-0320-8

O'Callaghan, F. V., Neumann, D. L., Jones, L., & Creed, P. A. (2017). The use of lecture recordings in higher education: A review of institutional, student, and lecturer issues. *Education and Information Technologies, 22*(1), 399–415. https://doi.org/10.1007/s10639-015-9451-z

Ogrinc, G., Batalden, P., & Moore, S. M. (2009). Realist evaluation as a framework for the assessment of teaching about the improvement of care/Commentary on "Realist evaluation as a framework for the assessment of teaching about the improvement of care." *Journal of Nursing Education, 48*(12), 661–667.

Owston, R., Lupshenyuk, D., & Wideman, H. (2011). Lecture capture in large undergraduate classes: Student perceptions and academic performance. *Internet and Higher Education, 14*(4), 262–268. https://doi.org/10.1016/j.iheduc.2011.05.006

Pawson, R., & Tilley, Nick. (1997). Realistic evaluation. London : Sage. EBSCOhost.

Richardson, A. M., Dunn, P. K., McDonald, C., & Oprescu, F. (2014). CRiSP: An instrument for assessing student perceptions of classroom response systems. *Journal of Science Education and Technology, 24*(4), 432–447. https://doi.org/10.1007/s10956-014-9528-2

Selwyn, N. (2014). Digital technology and the contemporary university: Degrees of digitization. Abingdon, UK: Routledge. Proquest Ebook Central.

Sorinola, O. O., Thistlethwaite, J., Davies, D., & Peile, E. (2015). Faculty development for educators: A realist evaluation. *Advances in Health Sciences Education, 20*(2), 385–401. https://doi.org/10.1007/s10459-014-9534-4

Stöhr, C., & Adawi, T. (2018). Flipped classroom research: From "black box" to "white box" evaluation. *Education Sciences, 8*(1), 22. https://doi.org/10.3390/educsci8010022

Taplin, R. H., Kerr, R., & Brown, A. M. (2014). Opportunity costs associated with the provision of student services: A case study of web-based lecture technology. *Higher Education, 68*(1), 15–28. https://doi.org/10.1007/s10734-013-9677-x

Whitworth, A. (2012). Invisible success: Problems with the grand technological innovation in higher education. *Computers & Education, 59*(1), 145–155. https://doi.org/10.1016/j.compedu.2011.09.023

Williams, A., Birch, E., & Hancock, P. (2012). The impact of online lecture recordings on student performance. *Australasian Journal of Educational Technology, 28*(2), 199–213. https://doi.org/10.14742/ajet.869

**Please cite as:** Bryant, M., Ryan, K., Simpson, C., Whitelock-Wainwright, A. & Wood, T. (2019). At the heart of a diverse technology: Applying a realist evaluation methodology to a university live streaming programme. In Y. W. Chew, K. M. Chan, and A. Alphonso (Eds.), *Personalised Learning. Diverse Goals. One Heart. ASCILITE 2019 Singapore* (pp. 59-67).