12-2016

Expectation-Centered Analytics for Instructors and Students

Suzanne L. Dazo
University of Nebraska at Omaha

Expectation-Centered Analytics for Instructors and Students

A Thesis

Presented to the

College of Information Science and Technology

and the

Faculty of the Graduate College

University of Nebraska

In Partial Fulfillment of the Requirements for the Degree

Master of Science in Computer Science

University of Nebraska at Omaha

by

Suzanne L. Dazo

December 2016

Supervisory Committee

Brian Dorn, Ph.D.

Neal Grandgenett, Ph.D.

Harvey Siy, Ph.D.

Abstract

Expectation-Centered Analytics for Instructors and Students

Suzanne L. Dazo , MS

University of Nebraska, 2016

Advisor: Brian Dorn, Ph.D.

Learning analytics is the measurement, collection, analysis, and reporting of data about learners and their contexts. An outcome and primary goal of learning analytics should be to inform instructors, who are primary stakeholders, so that they can make effective decisions in their courses. To support instructor inquiry, I apply theory on reflective practice to learning analytic development. Articulating an instructor's pedagogical expectations is one way to begin facilitating a reflective practice. Expectations based on instructor goals serve as a natural next step and the springboard from which data can be collected. I hypothesize that a learning analytic that encodes and reifies instructors' individual expectations will better support reflective practice for instructors and allow students to more reliably meet set expectations.

I took a user-centered approach to learning analytic research and development. First I triangulated empirical analysis of analytic use with focus groups to understand how instructors interacted with analytics. Instructors had a wide range of behaviors, needs and expectations. For most instructors, analytics were used very briefly (less than 1 minute). Instructors also requested a way to aggregate data from different analytics to better support their information needs. Based on these findings, I developed learning analytics within TrACE to allow for instructors to specify expectations and see student progress related to

those expectations. Students could also view their progress towards completing expectations.

Finally, I conducted a field study to compare both instructor analytic use and student compliance to expectations without and with the presence of these analytics. The results of the field study did not support the hypothesis. Instructors for the most part did not change their behaviors with the introduction of these analytics. Students also did not meet expectations more reliably, but one course saw a significant improvement in performance. Without visible expectations, students met significantly fewer posting expectations than other expectations. With explicit expectations, posting performance was no longer significantly less.

Acknowledgements

I would like offer my special thanks to my fellow CEL lab members whose efforts were invaluable in my research. To Nick Stepanek for his help with editing and data processing, to Aarjav Chauhan for working with me on the participatory design session and creating early mockups, and to Bobby Gibbs for helping test and tweak my prototypes, I am truly grateful.

To my good friends Susannah, Sammi, Johannes, Robert, and Nicole for cheering me on, always offering help when I need it, and reminding me that I had it in me all along. Thank you.

To my family, especially my mother Sharon and my brother Jonathan, I am very appreciative of your love and always supporting me in my endeavors.

Advice given by my committee members Dr. Harvey Siy and Dr. Neal Grandgenett have been particularly helpful in guiding me and keeping me on-track through this long process.

To Hunter Kennoy for all of the proofreading, motivational talks, and love, I'll always be thankful.

Finally, I would like to thank Dr. Brian Dorn for advising me through this process, challenging me to be a stronger researcher and writer, finding the time to meet with me, and always having a statistics book on hand.

This work is funded in part by the National Science Foundation under grant IIS-1318345. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

With the emergence and widespread use of technology for educational contexts, the use of learning analytics for classrooms has been praised as having made data about learners visible that has previously been "unseen, unnoticed, and therefore unactionable" (Bienkowski et al., 2012). In practice, this newly available information may not be as actionable as we expect. Instructors, who play an integral role in the classroom, are also necessary for the effectiveness of these analytics. Although data is more available, a major issue in facilitating interaction between instructors and students is that many existing learning analytics do not provide all of the information needed for instructor interventions to take place (Dyckhoff et al., 2013). This study aims to fill that gap through the development of a learning analytic designed towards instructor needs. Additionally, this analytic is evaluated by its impact on instructor and student behaviors in the classroom.

Learning analytics are defined as the 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 (Ferguson, 2012, Siemens and Long, 2011).

The goals of learning analytics are to improve course activities, identify problematic students/imbalances in the class dynamic, and allow for quick intervention by the instructor (Charleer et al., 2014).

Current research has given excellent evidence on what qualities can make for effective analytics (Scheffel et al., 2014) and many tools exist with the intention of informing instructors about student activities (see e.g., (Dawson et al., 2010, Romero et al., 2010)). However, Dyckhoff et al. (2013) identified that of almost 30 learning analytics analyzed, they do not yet answer all the questions teachers have for their classes. Additionally, instructors had challenges interpreting visualizations which are often created with the assumption that users are familiar with data mining techniques and complex analysis methods (Scheffel et al., 2011). Many learning analytics focus on reporting preset quantitative measures that may or may not be important to an instructor and do not cover the full extent of their data needs. In order for learning analytics to be more relevant to an instructor's practice, these analytics need to address the instructor's data needs for their specific course context. Learning analytic design should articulate the pedagogical intent of the teacher as opposed to being imposed by developers (Lockyer and Dawson, 2012, Wise, 2014).

Some setbacks include a gap in studies on the entire course context Greller and Drachsler (2012), a lack of time in development to involve teachers, and a lack of existing community that prioritizes the involvement of instructors in design (Nelson et al., 2008). It has been suggested that researchers should focus instead integrating learning analytics into everyday practice and develop better and more usable tools for learners and teachers (Chatti et al., 2012, Dyckhoff et al., 2013). In moving forward, one way to address these issues is with a *user-centered* approach that seeks to understand how instructors are actually using learning analytics for their pedagogical interventions in course settings. I do this with TrACE, a tool that has encountered many of these pitfalls in learning analytic design (Elson, 2016).

TrACE is an online video-playback platform that was developed with the goal of supporting collaboration among students within video-based course contexts (Dorn et al., 2015). Instructors upload videos to this platform for students to watch, and the system allows students to annotate and reply to annotations in videos. As a part of this system, an analytic dashboard was developed to allow for instructors to interpret student viewing behaviors, which was a need identified among several practitioners and education researchers (Maher et al., 2015, Lacher and Lewis, 2015). A study on this system in particular is useful as video elements are widely used in large-scale online education platforms (e.g. Coursera, Udacity, Khan Academy), but as a research prototype, it serves as a more flexible platform that is responsive to instructor needs on a smaller scale during development.

In developing learning analytics for TrACE and similar systems, it is necessary design based on education theory. The learning analytics cycle and reflective practice are two theories that can inform an improvement on the quality of intervention through learning analytics (more details in Chapter 2). The rationale behind this is that awareness and reflection support for educators are major goals for learning analytics (Scheffel et al., 2014). First, the learning analytic cycle (Clow, 2012) describes the flow of information in learning analytics. Some sort of intervention has an effect on learners, and these interventions are originally informed by learner-generated data in the form of metrics or analytics. Instructor actions are one type of intervention. Through analytics, instructors can make predictions about their students and take actions that would serve as an intervention on either current learners or future ones. To improve the quality of instructor interventions, the outcome and a primary goal of learning analytics should be to properly inform instructors so that they can make effective decisions. *Reflective practice* is a theory which can be a guide in how to aid instructor intervention. Reflective practice is a way in which instructors can consider the goals important to them in their course contexts, gather data, and process that data to accomplish or redefine those goals (Schön, 1987). When an instructor does not use reflective practice, he or she may not initiate any interventions or change his or her teaching strategies when students have issues which should be addressed (Sparks-Langer et al., 1990, Murphy and Ermeling, 2016). When analytics do not support reflection, instructors may be collecting data which overall does not support their inquiry. Although studies have been conducted to take into account instructor inquiry and the questions instructors want answered about their students, this work remains mostly in theory (Dyckhoff et al., 2013).

The first step in reflective practice is to have an instructor consider his or her goals. When an instructor has course goals, expectations based on those goals serve as a natural next step and the springboard from which data can be collected. In support of applying expectations to learning analytics, instructors have already expressed interest in being able to more quickly identify if expectations are met in TrACE (Elson, 2016). Other researchers have also made similar attempts at goal-based visualizations in learning analytics. Most notably, Muslim et al. (Muslim et al., 2016) utilized a workflow of eliciting instructor goals and questions to create visualizations that apply most to an instructor's needs. Additionally, making expectations explicit has been cited as a practice that is beneficial to students as well (Dennen et al., 2007).

A Learning Analytic that encodes and reifies instructors' individual expectations will better support reflective practice for instructors and allow students to more reliably meet those expectations. Currently, information is provided through analytics without directly taking into consideration an instructor's unique practice or what their goals and expectations may be (Schön, 1987, Van Manen, 1995). Reflective practice is only useful insofar as the information the instructor receives can help support or challenge his or her expectations. For instructors to effectively make observations that allow for reflection on their practice, learning analytics should present data directly related to their course expectations. As additional support to students, Sadler (1989) provides three conditions where students can benefit from feedback in academic settings. All of these rely on the transparency of course expectations and students understanding their own behaviors in relation to those expectations.

For reflective practice to take place, instructors should be aware of student behaviors in their class. To develop an analytic that supports this, the first phase of my research is a formative study with the goal of understanding current practice and the range of expectations instructors may have (Chapter 3). The results of this exploratory study will inform the design of my Learning Analytic. RQ1 and RQ2 are questions that I hope to answer through this initial exploratory study.

- **RQ1**-How do instructors currently conduct inquiry on student behaviors?
- **RQ2**-What expectations do instructors see as valuable to model within the context of learning analytics?

An expectation-centered analytic that translates the instructor's expectations and requirements for the course will be developed taking into account the results of the first phase of the research study (Chapter 4). This analytic will aggregate data relevant to instructors as opposed to instructors independently synthesizing conclusions from multiple sources, which can be difficult to interpret (Elson, 2016). The expectations specified by instructors will also be made available to students. By implementing an expectation-centered analytic, I can evaluate its effectiveness with regards to both supporting reflective practice for instructors and supporting students to answer several more research questions (Chapter 5):

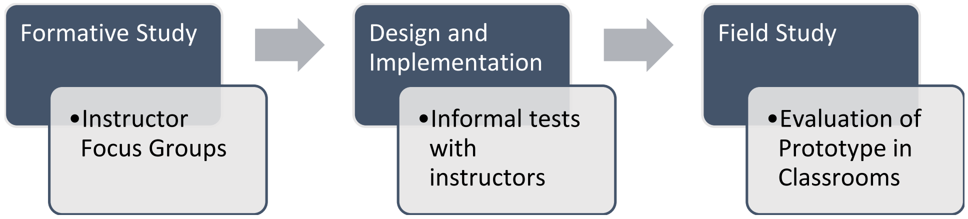


FIGURE 1.1: An overview of the format of the study

- **RQ3**-How does instructor inquiry change with the presence of this analytic?

- **RQ4**-How do student behaviors change with the explicit presence of this analytic?

To answer my research questions, I conducted a set of studies within one iteration of design-based research (Barab and Squire, 2004, Collins et al., 2004). Design-based research is an iterative methodology that allows for interventions such as learning analytics to be studied in the learning environment as opposed to a laboratory setting. This allows for immediate impact on the classroom as well as gaining insight through each iteration that can inform both theory and practice. My research was done within the context of a single iteration of the design cycle of TrACE. I added to TrACE by creating design alternatives for analytics available in the system. Following this, I initiated a field deployment to evaluate the impact of these alternatives in classrooms. Finally, I conclude with contributions to both researchers and practitioners, limitations, and direction for future work. An overview of the methods for this thesis is included in Figure 1.1 and detailed further in Table 1.1.

TABLE 1.1: Summary of activities

A learning analytic that encodes and reifies instructors' individual expectations will better support reflective practice for instructors and allow for students to more reliably meet said expectations.

Study Phase	Research Questions	Data collected	Analysis method	Outcome
Formative Study	How do instructors currently conduct inquiry on student behaviors? - Questions asked - Current workflow - Changes from reflection	-Focus groups/ interviews with instructors -Instructor activity data in analytics (frequency, time on task, analytics used in a "session")	-Descriptive Statistics - Kruskal-Wallis test to measure usage trends over time - Affinity diagram of transcripts	-Most instructors do not access analytics often or spend much time in them -Some analytics are more popular than others and better support inquiry -Workflow issues include frequent context switching, mental calculation, lack of aggregation/high-level views, lack of actionable data
Formative Study	What expectations do instructors see as valuable to model within the context of learning analytics?	Focus groups/ interviews with instructors	Affinity diagram of transcripts	- Expectations mostly related to viewing/posting - Expectations for viewing is as simple as "watch the video before class". Often implicit. -Posting expectations can be complex, as instructors want different kinds of collaboration - Expectations can change over time
Field Study	How does instructor inquiry and expectations change with the presence of this analytic?	-Instructors activity data in analytics - -(frequency, time on task, analytics used in a "session") -List of Expectations and related log data	-Within-subjects study - Descriptive Statistics -Wilcoxon signed-rank test and Fischer's exact test for Analytic usage trends over time - What expectations were made	-No sig. difference in overall duration or frequency between semesters - Visiting and time on task were still low for many -Some instructors shifted to visiting expectation analytics more often than other analytics -Many expectations were consistent throughout study period
Field Study	How do student behaviors change with the explicit presence of this analytic?	Student performance measured as $\frac{\text{completed amount}}{\text{required amount}}$ for watching, posting, and quiz-answering expectations	-Within-subjects study to compare behaviors with and w/o explicit expectations -Descriptive stats and Wilcoxon signed-rank test for performance over time in activity data	- Without explicit expectations, students were worse at meeting posting expectations than other expectation types in all cases - With explicit expectations, there was no longer a difference - Only one course had significantly higher compliance to expectations

Chapter 2

Background and Related Work

To better understand where learning analytics can be improved with regards to promoting instructor reflection and increasing student activity in the classroom, the theoretical underpinnings for reflective practice and an overview of existing Learning Analytic systems is presented in this chapter. I also cover previous work conducted in TrACE which also contributes to the motivation for this thesis.

2.1 Theoretical Background

Learning analytics as a field is a combination of several different disciplines including action research, education, and educational data mining (Chatti et al., 2012), but in order to improve the quality of learning analytics, I focus on the theoretical underpinnings behind the learning analytics process, notably reflective practice and the learning analytics cycle.

2.1.1 Reflective Practice

Reflective practice is the process through which professionals conduct inquiry on and adjust their own behaviors during practice (Schön, 1987). Although what defines reflection is still widely debated (Larrivee, 2008, Luik et al., 2014), I present Schön (1987)'s version of the reflective process as an active and iterative process consisting of the following stages:

1. Data collection- Observations on the current situation involving spontaneous and routinized responses.

2. Surprise- The practitioner experiences an unexpected outcome from the data collection process that challenges their current knowledge.

3. Reflection- The practitioner considers the unexpected event as well as their current knowledge that led up to it.

4. Question structure- reflection on the thought processes that led up to this unexpected situation. Restructuring strategies of action, understanding, or framing of the problem.

5. Experiment- Take new actions and explore the newly observed phenomena. This could yield the hypothesized results or lead to more unexpected outcomes and thus more reflection-in-action.

Reflective practice is not attempting to find standard solutions to any given problem, but uncovering the details of the problem through gradual discovery which eventually leads to interventions. It is a cyclical process that iterates between theorizing about the current situation and experimentation, intervention, and observation of the situation.

Traditional experiments involving hypothesis testing are controlled. In contrast, reflective practice is a combination of exploration, move-testing, and hypothesis testing (Schön, 1987). That is, the practitioner may intervene only to see what happens, may influence the situation with an intended outcome in mind, or simply observe if the outcome matches a predefined hypothesis. While the practitioner shapes the situation through interventions, s/he remains open to the possibility that these interventions continue to produce unexpected outcomes. Using TrACE as an example, an instructor On the other hand, a failure in the reflective process occurs when the practitioner tries to completely control the situation. A failed reflective process includes set tasks where all input works solely towards that task. The instructor filters out evidence that could have led to reflection or changes in the class in favor of reaching predefined goals. Additionally, the practitioner may avoid being “wrong” and does not share information to other parties (such as students) that may influence the situation.

Let us consider two examples that might happen in using learning analytics. Two instructors believe that students who watch a video will learn more and do better on assessments. The first instructor wants students to watch the video and enforces this with participation points. In spite of students “watching” the content, test scores do not improve. The instructor continues to believe that students are not watching enough, and requires a higher percent of the video watched and assigns more participation points. The second instructor initially had the same requirement and also saw poor assessment results in her class. She uses the analytics and realizes that many students are watching the content, but many are fast-forwarding through the video. This causes the instructor to reconsider watching alone as a goal, and changed her goals to focus on comprehension instead. She adds automatic pauses and reflection prompts throughout future videos as an experiment to see if students will slow down and more closely consider the course content as they watch. The first instructor was not reflective and did not stop to reconsider if watching was the right expectation to have for students. The second instructor noticed a surprising trend, reflected, and adjusted practice accordingly.

Two scenarios, one in which the instructor fails to utilize the reflective process and one in which the instructor embraces it, are described as single and double-loop learning by Argyris and Schön (1978). Single loop learning encompasses the behaviors wherein a practitioner focuses on achieving a set goal in the most efficient way. The goal is perceived as immutable and no reflection takes place. Either the goal is met or it is not. Conversely, double-loop learning occurs when the results of an initial action leads to questioning those initial goals and values and revising the underlying assumptions that started those actions in the first place. It is within double-loop learning that reflective practice takes place.

Reflective practice has been studied by many scholars in an attempt to characterize these different levels of reflection and create effective measures of reflective practice (Larrivee, 2008, Jay and Johnson, 2002, Sparks-Langer et al., 1990). Larivee defines four levels of reflection in practitioners:

1. Pre-reflection - no active reflection and the instructor does not adapt their own teaching based on the students responses and needs.
2. Surface reflection - An instructor's strategies work towards a predefined goal, the instructor focuses on "what works" instead of considering instructional value of their goals.
3. Pedagogical reflection - The instructor reflects on their educational goals, the theories behind their approaches, and connects between theory and practice.
4. Critical reflection - An instructor considers the moral implications of their practice and reflects on their own beliefs and how it affects their expectations and teaching.

These levels of reflection are present in other works as well (Jay and Johnson, 2002, Sparks-Langer et al., 1990) although often titled-differently (surface reflection to descriptive/initial understanding and pre-reflection to habitual actions) and have been validated in each researchers' own educational contexts.

When applied to learning analytics, not only could an instructor use an analytic to reach his/her initially desired outcome, but through reflective practice, challenge and redefine those initial goals (Clow, 2012). Although reflective practice occurs personally and *in situ*, changes in expectations, actions, and goals could be external indicators of this process. Poor reflective practice would be observed as an instructor using the analytics for a fixed purpose that does not change throughout the semester, and not sharing or intervening with students based on the results discovered in analytics. In the next section, I discuss the learning analytics cycle which applies the theory and process of reflection-in-action to the context of learning analytics.

2.1.2 The Learning Analytic Cycle

The learning analytics cycle is an iterative process that is used to engage learners in their educational environment. Reviewing it once again (Figure 2.1), I discuss the different elements of the process:

Learners can be students studying in a course or participants in informal education. Learning analytics both starts with and should affect learners.

Data can be about learners or generated by them. Examples include data traces such as demographics, posts, test results, and click-level activity data (e.g. interactions within TrACE or Blackboard). This data needs to be processed and interpreted.

Metrics/Analytics provide insight into the learning process. These can include traditional dashboards, visualizations, or identifying specific students based on the data. These metrics inform the next step of the cycle.

Interventions have some effect on the learners. Interventions can include a dashboard for the learners to reflect on their own actions, or take place when an instructor directly addresses high-risk students. It should be noted that intervention does not have to occur with the same group of students. Data from one semester could lead to an intervention in a new semester.

The data collected from the Learning Analytic Cycle may not be uniform or may come from multiple sources. It is imperative to the success of a learning analytic that the data is pre-processed (cleaned, integrated, transformed, etc.) before being presented in the metrics phase (Chatti et al., 2012). The goal of the metrics phase is to provide insight through previously unobservable patterns. In this way, the metrics supports the data collection phase in reflective practice, and also allows for the practitioner to more easily notice unexpected outcomes to initiate the reflective process. It is during this intervention phase of the learning analytics cycle that reflective practitioners (Schön, 1987) reflect on their practice. Instructors can self-reflect on the effectiveness of their learning or teaching practice based on the results discovered in the analytics.

While not mentioned in the figure but mentioned in Chatti's learning analytics process (Chatti et al., 2012), an additional post-processing phase is involved. This allows for constant improvement of the analytic process. This could involve collecting new data, refining the data, or looking at new analytics altogether. This ties back to reflective practice, as ideally an instructor should be able to manipulate the data available to them based on their new goals.

The Learning Analytic Cycle is a model for learning analytics that draws from education theory. The key step is ensuring that information generated from learning analytics feeds back into learners via interventions. All design and development should be done with

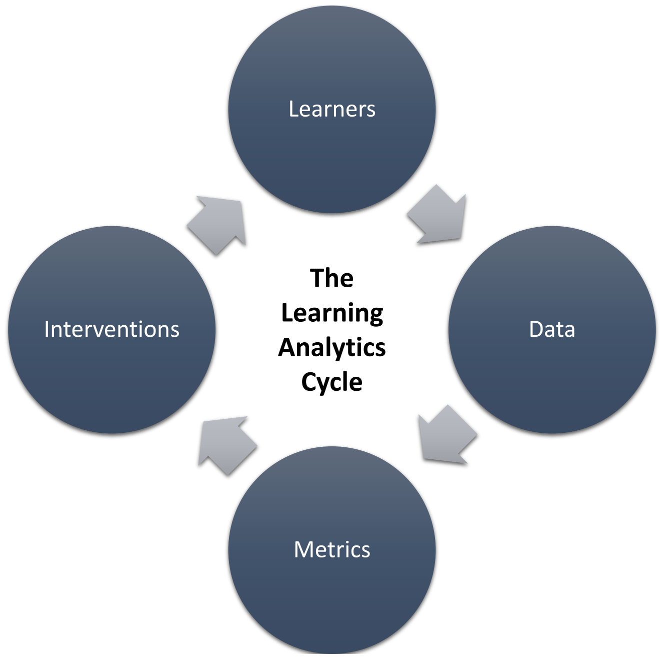


Figure 2.1: The Learning Analytics Cycle

this model in mind. Clow (2012) claims the Learning Analytic Cycle "instantiates and enables reflective learning" in the intervention step, and understanding reflective practice can help us determine if these interventions do support instructor reflection. In theory, metrics(analytics) should support reflective practice by allowing for quick observations related to course goals. These metrics should also be presented in a way that can bring attention to data that challenges those existing goals. This, in turn, should lead to reflection and changes in practice.

2.2 Implementation of Systems

Instructors and students are both primary stakeholders in learning analytics. However, most studies involving different stakeholder groups in learning analytics research target intelligent tutoring systems or researchers (78%) instead of students(12%) and teachers(18%) (Chatti et al., 2012). This research fills this gap in the literature to address the needs of students and teachers for learning analytics. Usually, students generate the traces that become the data presented in an analytic dashboard, and the teacher (should) analyze this information to inform his or her practice. Thus, it is important to consider how learning analytics have been implemented for both instructors and students, and the research that has been performed for both groups.

2.2.1 Instructor-centered analytics

Guidelines for instructor-centered analytics emphasize that instructors should be considered in the design of learning analytics (Scheffel et al., 2014). The information most important for teachers include overall success rate, mastery of concepts, frequent mistakes, and support for self-awareness and reflection (Scheffel et al., 2011). However, the current tools do not yet answer all the questions that teachers have in regards to the educational setting in which they are situated (Dyckhoff et al., 2013). In reality, current analytics are effective at answering questions about quantitative measures of use (*what is the student doing?*), but do not collect more complex information. More complex information can include qualitative evaluation(*does the student like the system?*), differentiation between groups of students, differentiation between learning offerings (*is the student choosing online offerings instead of in-class?*), data consolidation/correlation (*What percent of the learning modules are students using?*), and effects on performance (Dyckhoff et al., 2013). Other analytics that are not available include information about the instructor's own actions or information from multiple data sources (Dyckhoff et al., 2013).

Some studies have looked into the qualities of effective learning analytics (Scheffel et al., 2014), and others have researched what instructors might want from learning analytics (Dyckhoff et al., 2013). Often, the target population of these studies have been Learning

Analytic researchers or instructors who have no prior experience with learning analytics. A challenge here is that it is difficult for participants to imagine features and tools they would like without having any prior experience (Gulliksen et al., 1999). Collecting data from these participants produces results that focus more on data collection, privacy, and acceptance of a learning analytic system (Scheffel et al., 2014) as opposed to instructor desires to better understand the learning process of students (Goodyear, 2010). Interviewing instructors who are already experienced users of learning analytics may yield more fruitful results. By collecting information from experienced users, we might gain more insight into the challenges these instructors face in practice that may not have been considered by researchers working outside of the classroom.

2.2.2 Student-centered analytics

Student-centered analytics allow for students to have self-directed and self-regulated learning. Students should regulate their own performance in order to meet the goals and expectations of a course. Nicol and Macfarlane-Dick (2006) cite 7 ways that self-regulation can be supported in students:

- Clarify what good performance is. Students can only achieve goals if they know what these goals are in the first place.
- Facilitate self-assessment
- Give feedback information in relation to goals. Assist students in taking actions to bring themselves closer to accomplishing these goals.
- Encourage teacher and peer dialogue
- Encourage positive motivation

- Provide opportunities to close the gap between their performance and success. Allow at-risk students to understand their behaviors and correct them as necessary.
- Use feedback to inform teaching, such as with reflective practice

Additionally, for students to benefit from feedback, they need to understand what good performance is for the course, how their current performance relates to ideal performance, and how to act to close this gap between their current performance and good performance. This emphasizes the need for making the expectations of instructors available to students through learning analytics. Learning analytics are not solely for the instructor or solely for the student, and making expectations clear can benefit both students and instructors in improving achievement in courses.

Several existing analytics attempt to provide students with information for self-reflection of their learning. Signals, a Learning Analytic from Purdue University, (Arnold, 2010, Arnold and Pistilli, 2012) manually collected student use data from a Learning Management System (LMS) and provided feedback on progress using a stoplight system. The goal of Signals is similar to mine: to provide analytics with actionable feedback. The presence of this information allowed students to make corrections as they realized they were off-track within the course, and students with this intervention sought help earlier. Faculty also saw that students were more proactive. Students expressed a desire for more specific information to how on-track they were, and instructors desired more action-oriented and helpful feedback beyond a good (green)/fair (yellow)/poor (red) metric.

Duval (2011) analyzed various learning analytics for students and emphasized that visualizations in relation to a goal can be more effective than by being presented as raw data. However, no explicit Learning Analytic examples were provided, although systems such as health trackers can be used as guidelines for goal-based analytics. While many works focus on either student analytics or instructor interventions, none found talk about the intersection between the two, and even large scale literature reviews fail to find current systems that allow this kind of interaction (Dyckhoff et al., 2013). Overall, while students are provided with self-regulating information on their own actions, instructors are left with either basic information that does not truly inform their practice, or they are left out altogether by not being involved in the analytic process.

My contribution to this body of knowledge aims to "bridge the gap" through application and design guidelines between learners and instructors that use analytics. This motivates the creation of an analytic that allows for both the instructor to convey what they want to know from the student, and for the student to understand how their behaviors match with instructor expectations. Additionally, I aim to create an analytic that is informed by the learning analytic cycle model, which claims to support instructor reflection when followed (Clow, 2012). Finally, this research presents a unique opportunity to evaluate analytics both in the classroom context and with experienced instructors.

2.3 Preliminary Research

Preliminary research in TrACE informs and motivates the work of this thesis. A qualitative study consisting primarily of a thematic analysis of instructor journals and instructor interviews was conducted by Elson (2016) to gain insight into instructor formative assessment practices. Several themes (*Knowledge of Students*, *Actions*, and *Limitations/Shortcomings*) and subcategories (*Student Behavior Relative to Expectations*) related closely to the work proposed in this thesis. To elaborate further, some of these categories are expanded on with qualitative examples.

First, *Student Behavior Relative to Expectations* was categorized as a subsection of the *Knowledge of Students* theme. This section as described by Elson showed the ways that TrACE enabled awareness of student performance relative to instructor expectations. Important factors here were the ability to quickly assess if assessments were met and being able to assess the class as a whole as well as individual students. These expectations covered the range of actions in TrACE (watching, posting, etc.) but instructors were very interested in knowing if students are meeting these expectations or falling short.

The *Educator Action* theme was described as the motivations behind instructional change or intervention based on the insights/data presented to instructors through TrACE. Tying back to reflective practice or the learning analytic cycle, this would be the intervention that takes place as a result of evaluating if an instructor's goals were being met. Just as in the other related work, Elson noted that these interventions could be with a single student, with the whole class, or with the next iteration of a course.

The final theme from Elson's work that related to this thesis was the Limitations and Shortcomings presented by instructors with regards to the system. Educators advocated for system features to be available to students. Two of the six educators interviewed mentioned a desire for analytics students could view to help them know if they are doing what is expected of them. Elson proposed that such analytics would directly benefit instructors by encouraging students to interact with the system in a way that better meets instructor goals.

From this related work, there is an opportunity to create analytics that support reflective practice for instructors. Also, specifying expectations and an expectation-centered analytic is something that instructors have expressed a desire to have in order to improve their practice. In the following chapter, I present a formative study that directly informs the development of expectation-centered analytics.

Chapter 3

Formative Study

In order to assert the hypothesis *A learning analytic that encodes and reifies instructors' individual expectations will* ***(1)****better support reflective practice for instructors and* ***(2)****allow students to more reliably meet said expectations*, a three-phase study was conducted. A formative/exploratory study is the focus of this chapter, and serves as a form of requirements gathering in which the results inform the design of an expectation centered analytic around instructor expectations. Future chapters build upon the study in this chapter by taking findings and developing an analytic prototype (Chapter 4) and evaluating said prototype (Chapter 5).

The goal of this formative study was twofold: *(i)* to offer some insight into some of the expectations instructors might have for their students and *(ii)* the extent to which analytics were meeting their needs in order to establish design guidelines for analytic development. To do so, I triangulated quantitative activity-log data of instructor use of TrACE from previous semesters along with data collected from a participatory design session (Kensing and Blomberg, 1998) involving instructors.

In order to set the context for this study, it is important to go into further detail on

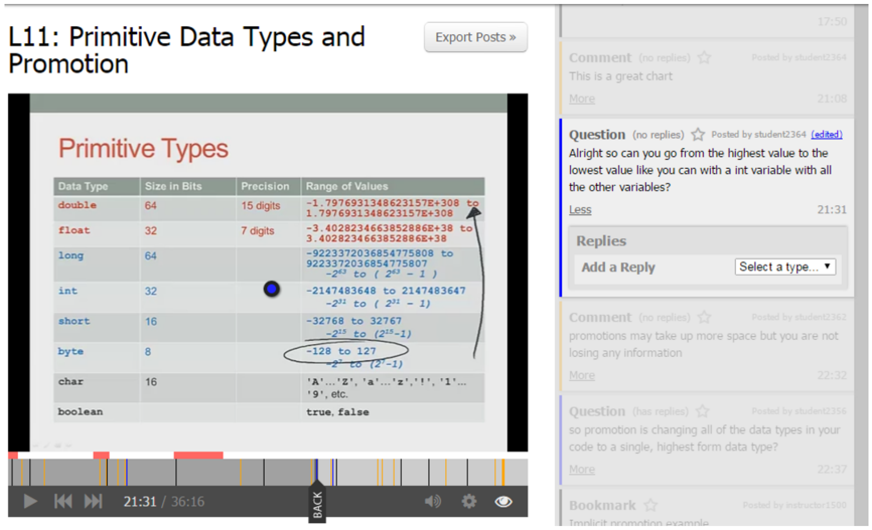
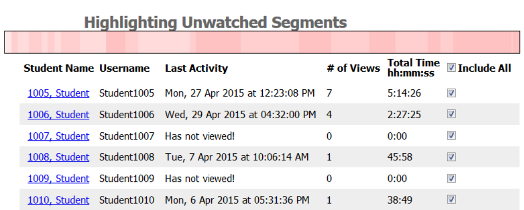


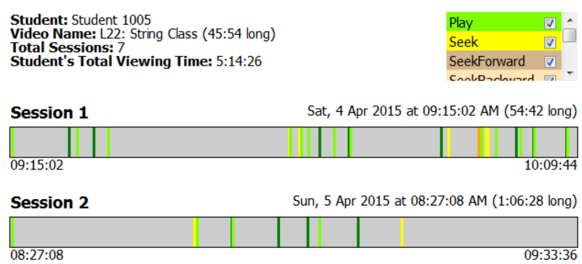
Figure 3.1: The video playback system in TrACE. Students can create annotations pointing to specific video content

how TrACE works and how instructors and students may use it. Instructors upload videos to this platform for students to watch, and the system allows students to annotate and reply to annotations in videos (Figure 3.1).

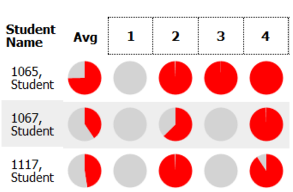
As a part of this system, a collection of learning analytics were developed to allow for instructors to interpret student viewing behaviors. There are 8 analytics overall within



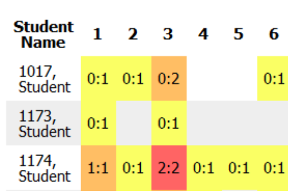
(a) Media Activity



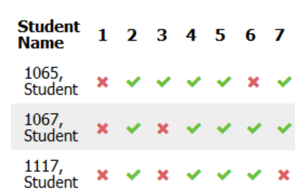
(b) Session Summaries



(c) Percentage Viewed



(d) Annotation Summary



(e) Viewing Summary

Figure 3.2: Sample images of the analytics available to instructors in TrACE

the system: Media Activity, Session Summary, Annotation Summary, Loyalty, Recency, Percentage Viewed, Viewing Summary, and View Count graph. Most of these analytics are at the video level for a single course and show activity for all students in that class unless otherwise specified. To better understand these analytics, I will briefly describe what questions instructors can answer and the visualization used. Screenshots of a subset of analytics are also provided (Figure 3.2).

Media Activity (Figure 3.2a)- For a single video it answers how many times, what parts, and when did students watch a video. It also shows how much time students spent and what actions they took while watching. These answers are summarized as a complex presentation of timestamps, view count numbers, and a video playback bar that fills in what portions of the video were watched in aggregate by students.

Session Summary (Figure 3.2b) (accessed via Media Activity)- For a video it answers similar questions to Media Activity (how long, how often, what actions did they take) but focused on a single student. It is presented as a list of sessions (from opening the video to leaving the video page) where each session contains a timeline with markers indicating what actions occurred.

Percentage Viewed (Figure 3.2c)- A pie chart showing what % of the video students watched.

Annotation Summary (Figure 3.2d)- Answers how many posts and replies did students make presented as a heatmap.

Loyalty - Answers how many times did students open a video also presented as a heatmap

Recency - Answers when was the last time students interacted with the video also presented as a heatmap.

Viewing Summary (Figure 3.2e)- A check (yes) or X (no) answering if students opened the video.

View Count graph - A histogram showing how the number of views differ (between students, between videos, over the semester).

In order to understand instructor needs in both assessing students and reflecting on their own practice, I addressed the following research questions, outlined below:

- **RQ1**-How do instructors currently conduct inquiry on student behaviors?
- **RQ2**-What expectations do instructors see as valuable to model within the context of learning analytics?

TABLE 3.1: Instructors using TrACE from 2015 to present

Instructor	Institution	Subject	# Courses Taught Sp15	Fa15	Sp16
1	A	Calculus 1 &2	3	2	2
2	B	Information Assurance	1	1	1
3	A	Calculus 1	2		2
4	B	Intro to Computer Science	3	3	3
5	A	Education	2	2	2
6	A	Scientific Inquiry	3		9*
7	C	Business Intelligence	2	3	3
8	A	Calculus 1	1		
9	B	User Interfaces and Design	1	2	
10	B	Intro to Computer Science			1
11	B	Intro to Computer Science			1
12	A	Political Science		1	
13	B	Database Administration		1	
14	A	Scientific Inquiry		2	

3.1 Methods Overview

The participants for this study included 14 instructors that used TrACE's analytics from January 2015 to May 2016 (Spring '15 to Spring '16 in North American vernacular). Excluding one instructor who did not use analytics, Table 3.1 includes basic information about the instructors, including how many courses they taught in the Spring '15, Fall '15, and Spring '16 semesters. Instructors were assigned a random ID and course names were generalized for anonymity. The majority of these instructors had used TrACE prior to this study, so they may have already formed habits in their analytic use. This was a key distinction from many other studies on system use, which were tested with first-time users of learning analytics (Arnold and Pistilli, 2012, Muslim et al., 2016, Ali et al., 2012). These classes were small to medium in size, with the largest class having 59 students. There was a mix of undergraduate and graduate courses, the majority being in STEM disciplines with a few education and political science courses as well. On average, instructors taught 1.8 courses per semester with a maximum of 3 courses in any given semester (one instructor had 9 "courses" in the system, but this was a single class divided into groups). Some instructors taught multiple sections or taught the same course across multiple semesters with the same video content. During the study, instructors were given free reign over how the analytics were used, and they were only provided with an introductory tutorial on the analytics at the start of each term.

Fine grain data was collected on instructor analytic use within TrACE. This enabled me to analyze how instructors were using the system at the time, which analytics were preferred, and some basic patterns of behavior (having switched between many analytic screens or only looking at one). These methods are elaborated upon in Section 3.2. Although insights into instructor behavior were difficult to infer from the data alone, combining this with focus group data provided a clearer picture of the context in which instructors conducted inquiry on their students.

The qualitative portion of this formative study included two 2-hour participatory design (Kensing and Blomberg, 1998) sessions where instructors were invited to discuss their inquiry process and brainstorm analytic designs that would help support those inquiries. The first part of the design sessions was a focus group related to their expectations in courses that used TrACE, and the second consisted of a brainstorming and sketching session where instructors produced analytic designs. The methods for the qualitative portion are elaborated upon in Section 3.3.

3.2 Activity Data in TrACE

To answer my first research question, *How do instructors currently conduct inquiry on student behaviors?*, I analyzed the data from 14 instructors to find out:

- How often do instructors visit analytics?
- How long are instructors spending in analytics?
- Do instructors prefer some analytics over others?
- Is interaction consistent across semesters?

Logfile data that informed this study included which analytics were accessed, who accessed them, timestamps, and other action details (such as applying filters or closing reports).

To answer these first two sub-questions, the frequency and duration of analytic use were calculated for each instructor. An action was logged every time an instructor entered an analytic, took an intermediate action (e.g., changing filters, looking at different students, changing the video targeted for analysis), exited. If a session timed out (i.e. there were no consecutive actions for at least 15 minutes) then the duration of that session was calculated using the timestamp of the last action recorded. Sessions with durations of less than one second were filtered out, as they were likely misclicks where the instructor would not have gained any useful information from the analytic. Duration data was not evenly distributed among instructors, so medians and non-parametric tests were used for my analyses.

To understand instructor changes in behavior, instructor analytic visit frequency was directly compared across semesters both in a raw form and as a ratio of Frequency/# Videos in all courses. I used a Visits to Videos ratio because an instructor may not have been teaching the same courses or using the same videos every semester. Finally, the frequency of visits was also split between each analytic in the system, and calculated as the proportion of total visits. For the duration of instructor visits, instructor data was not normally distributed, so a Kruskal-Wallis test was used to answer the last two sub-questions.

3.2.1 Results

First, I wanted to answer how often instructors visited analytics through an analysis of the frequency of visits. The instructors used the analytics within TrACE 1268 times overall with a distribution of 494 sessions in Spring '15 (39%), 410 sessions in Fall '15 (32.3%), and 364 sessions in Spring '16 (28.7%). How often instructors visit the analytics was not evenly distributed for any of the semesters observed. Table 3.2 presents how often an instructor visited any analytic normalized by the number of overall videos in their course. Instructors that visited more than once per video are in bold. 69.3% of instructors did not visit an analytic at least once per video, so the majority of instructors may be viewing multiple videos for each visit, or not viewing analytics for those videos at all.

TABLE 3.2: Frequency of overall analytic use normalized by number of videos in all courses

	Sp15	Fa15	Sp16
Instr 1	0.35	0.67	0.33
Instr 2	**10.3**	**5.3**	**4.4**
Instr 3	0.22		0.22
Instr 5	**1.27**	**2.56**	**1.03**
Instr 6	**1.97**	**1.37**	0.61
Instr 7	0.21		0.02
Instr 8	0.26	0.14	0.14
Instr 9	0.74		
Instr 10	0.27	0.81	
Instr 11			0.29
Instr 12			**1.5**
Instr 14		0.77	
Instr 15		0.09	

TABLE 3.3: Median duration (in seconds) instructors spent in all analytics. Instructors with a significant difference in duration ($p < 0.05$) are in bold

	Sp15	Fa15	Sp16
Instr 1	6	18.5	42.5
Instr 2	19	17	20
Instr 3	52		24
Instr 5	42.5	61	58
Instr 6	35	28.5	83.5
Instr 7	31		35
Instr 8	8.5	33	38
Instr 9	11		
Instr 10	35	44	
Instr 11			35.5
Instr 12			70
Instr 14		27	

Continuing to the next subquestion, I looked to then answer *how much time do instructors spend in analytics?* Table 3.3 shows the median time instructors spent in the analytics overall. This median was around 30 seconds overall for instructors, but this greatly varied (6 seconds up to 83.5 seconds). Regardless of the variation, it was apparent that most instructors did not spend much time using analytics, and only three instructors had medians greater than one minute. Comparing frequency to duration, Instructors 4 and 11 stood out for having spent more time in the analytics and also visiting at least once per video. These

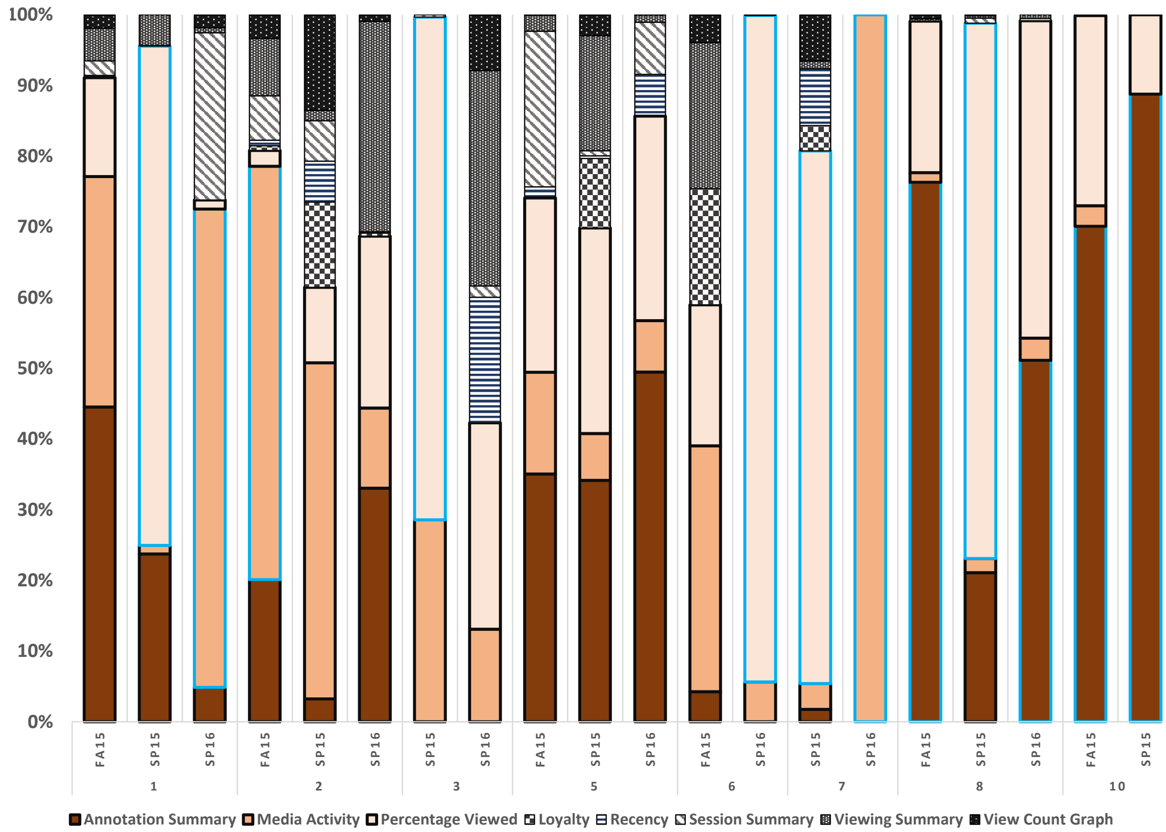


Figure 3.3: Proportion of analytic use by instructor, compared across semesters. Excludes instructors who only used TrACE for one semester.

users may have have developed a consistent procedure to analyze student behaviors in every video.

I next looked at whether instructors demonstrated a preference for some analytics over others. Five instructors in Spring '15 (55.6%), one instructor in Fall '15 (12.5%), and four instructors in Spring '16 (44.4%) had the majority of their sessions in a single analytic. Combined, a third of all instructors had a majority analytic. Media Activity, Percentage Viewed, and Annotation Summary were the most popular analytics (Figure 3.3).

Preference can also be broken down by duration to determine if instructors also spend more time with some analytics over others. When performing a Kruskal-Wallis test to compare differences in duration between analytics within a semester, most instructors did not have a significant difference in duration between analytics. This could be either because there were not many visits to the analytics individually in the first place, or there was no preference at all. Excluding two instructors that only use a single analytic, only 4/11 instructors showed a significant preference for an analytic in at least one semester. These 4 instructors were instructor 2 ($X^2(7) = 41.1, p < 0.001$) in Spring '15, instructor 3 ($X^2(2) = 9.681, p = 0.008$) in Spring '15, instructor 4 in Fall '15 ($X^2(7) = 34.3, p < 0.001$) and Spring '16 ($X^2(7) = 14.67, p = 0.04$), and instructor 11 ($X^2(6) = 15.7, p = 0.015$) in Spring '16. To determine popular analytics (by duration, instead of frequency) I ranked the analytics by the mean usage (as duration) and calculated how many instructors spent the most time in that analytic. The results show that Media Activity (5), Percentage Viewed (3), Recency (2), Annotation Summary (2), and View Count Graph were used the longest. Media Activity and Percentage Viewed were also the most frequent, so some instructors were both dedicating more time and visits to these analytics.

To answer the final sub-question and understand if there were any differences between semesters, I compared frequency of analytic visits (visits per video) as well as duration (Kruskal-Wallis test). Instructors who visited often (at least once per video) continued to do so. Only one instructor who visited often decreased to less than one visit per video in a future semester. Instructors who did not view analytics often also maintained their trends; no instructor that viewed less than once per video ever changed their habits to visit more often than once per video. Another notable trend among all instructors who used analytics for multiple semesters was that from the Fall '15 to Spring '15 term, every instructor had visited the analytics either the same amount or less per video. This could be an indicator that instructors have plateaued in their use of analytics.

When comparing an instructor's duration across multiple semesters, half showed a significant ($p < 0.05$) difference in access (Instructors 1, 4, 5, 7). For instructors who spent more time in the analytics but visited less often, this could mean these instructors were analyzing more courses and videos in one sitting as opposed to shorter more surface level bursts. On the other hand, it could indicate that the analytics were complex and difficult to parse so more time was needed to understand them. These possibilities were kept in mind when interviewing instructors about their analytic use (Section 3.3).

It was becoming more clear that Percentage Viewed and Media Activity were commonly used by many instructors. In looking more closely at these specific analytics, I propose some possibilities as to why these analytics were so popular. Initially, the Viewing Summary and Percentage Viewed analytic seem to answer the same question (Did my students watch the video?). When directly comparing the two, Viewing Summary can be misinterpreted as a false positive, as just opening a video counts as a "watch" whereas Percentage Viewed is much more detailed and allows for the instructor to see how much content was viewed at a glance. Media Activity is a much more detailed analytic that can answer a variety of questions for instructors (i.e. When are my students watching?, How often are they watching?, What parts have they watched?). Two of these questions: when (Recency) and how often (Loyalty) are covered in other analytics that were not used as often by instructors. Media Activity allows instructors to find both pieces of information in one location.

Annotation Summary was another common analytic, and it is the only analytic designed to report on student posting behaviors. This makes it the only option available for instructors who want to know if students were participating without opening videos and reading individual posts from students. Which analytics were most used could also tie into expectations. If learning analytics support instructor intervention as related to their goals,

Table 3.4: Participatory Design participants

Instructor	Institution	Number of Courses	# Semesters Experience
Inst A1	A	1	1
Inst A2	A	3	2+
Inst A3	A	1	2+
Inst B4	B	1	1
Inst B5	B	2	2+
Inst B6	B	1	2+

certain analytics may be tied to specific expectations of students (watching, posting, or otherwise). While most instructors did not have a clear preference for an analytic, these more commonly used analytics could hint at the questions that instructors were most interested in answering about their students and the types of expectations that instructors had. The following qualitative study explored these questions and instructor needs more in-depth.

3.3 Qualitative Study

In Spring 2016, two 2-hour participatory design sessions were conducted with three instructors each at Institution A (A metropolitan doctoral university in the Midwest) and Institution B (a medium-sized residential private university in New England)(Table 3.4). All instructors who had used TrACE over the past calendar year were invited via email to participate and were compensated for their time with Amazon gift cards. Of these 13 instructors, 6 total instructors had accepted the invitation. Four of the six instructors had also taught the courses from the Fall 2015 student surveys reported on earlier. A description of the methodology for this participatory design session will be split between the focus group and the rest of the design session.

3.3.1 Focus Groups

The focus group was 25 minutes and consisted of 4 questions related to their expectations:

Instructor Expectations-What expectations do you have for your students, and how do these expectations relate to your overall goals for the course?

Evaluation of Expectations-To what extent are you aware that students are meeting the expectations? How confident are you that your students are meeting this expectation?

Clarity of Expectations-To what extent do you enforce these expectations of students? How do students know what the expectations are?

Changes in Expectations-How have your expectations changed since you've started using TrACE?

The focus group was transcribed and analyzed to extract examples of expectations, goals motivating these expectations, and how these expectations were situated within the classroom context. Results of these focus groups were transcribed and analyzed to expose the range of expectations that instructors may have for their students. Affinity diagramming (Beyer and Holtzblatt, 1999) is a form of contextual inquiry through which work activity notes, or details on the instructor's current workflow were extracted and grouped. These groupings were then labeled and used to better understand the general needs, problems, functional requirements and nonfunctional requirements that the system needs to fulfill for these users.

3.3.2 Participatory Design Session

Following the focus group, instructors were shown aggregate data represented as graphs of the Spring 2015 data. The two graphs included data on frequency over the course of a semester, and the duration in the various analytics of TrACE. They were asked about their initial interpretations of the data as well as how this reported data relates to or contrasts their current use of TrACE. Transcripts from this portion of the session were analyzed to extract instructor descriptions of their current use of the system, hypotheses for past system use, and rationale for their current system use. Comments such as “We’re getting tired at the end of the semester. There’s a lot of stuff going on” and “There are a little bit complex metrics and sometimes I don’t have the time to process what you’re telling me. So I tend to use them very infrequently” were examples of what statements were extracted.

Next, instructors were asked to report on questions that they have about student behavior. The questions asked to prompt instructors are included below:

- What are some questions you have thought about students in the class you’re teaching that uses TrACE?
- As a teacher using TrACE, how does the answer to this question help you?

To answer the first prompt, each of the instructors came up with a list of questions individually. Then, they collaborated and reduced the list to 6 questions which they thought were of top priority and that encapsulated most of their areas of concerns. The second prompt was then shown to instructors. Individual responses were placed next to the corresponding question. Instructors selected their top priority questions and, in groups, they were asked to sketch out what an analytic or visual aid would look like that would help answer a given question. The instructors worked with the researchers/developers in sketching out their ideas on paper. There were three smaller sketching sessions that lasted 15 minutes, for a total of 45 minutes. After the sketching session was complete, instructors explained their sketches to the other participants and researchers/developers for another 20 minutes.

Artifacts from the participatory design session included the exhaustive list of questions and rationales generated by instructors and the sketches created as a result of the design session. Partial transcripts from audio recordings documented the experience. During sketching sessions and when conversations overlapped, transcripts were supplemented by field notes from the three researchers present.

3.3.3 Results

Table 3.5 is the exhaustive list of the questions that instructors asked along with how it would be useful for them. One question was excluded due to the fact that it was not posted as a question, instructors could not provide reasons for how it would be useful to them, and it was not selected for sketching.

In the first participatory design session, instructors sketched either alone or in a pair, with pairs rotating for each sketch. This produced 5 designs in total covering 5 questions. One group had reworked an existing sketch a second time, so there was one fewer sketch than the intended six. In the second design session, only a single SMART board was available for sketching. The SMART board could accommodate two participants at once, so all three participants worked together to design analytics from three questions. Although all of the designs were analyzed, a sample of the designs are presented here. These samples were chosen because they embodied many of the common responses and needs instructors had throughout this study. There were several questions posed that related to improving the quality of their course and supporting their goals (such as developing a community of learners), but did not easily translate to an analytic that an instructor could use.

Some of the analytics sketched by instructors focused on the student view and new functionality within TrACE as opposed to a visualization the instructor could use in an analytic. One example of this was a design answering the question "What would get the students more engaged to the content and community?"(Figure 3.4). The rationale behind this was that there was a need for students to know their status on videos in order to get

Questions	How it can be helpful
Is this lecture useful/engaging?	Feedback over time and semesters on video content Let's me know if the lecture should be changed to increase comprehension Allow me to rework lectures or portions of lectures to better serve students
Are students engaged when watching the video?	Students need to pay attention to do well Know if students are exhibiting passive vs. active learning See that students are committed to what I'm trying to teach Where I need to change vs. where the student needs to change
Are my students confident in the subject matter?	Feedback on student performance Help identify students that may need additional help with material Allow me to guage proficiency but also how solidly they believe in their knowledge
What would get the students more engaged to the content and community?	ID barriers to community participation ID barriers to understanding video content Help alter delivery of content such that it is more meaningful and interesting to students Tailor in-class activities or online interactions
What are barriers to them asking questions and how to address them	Be able to encourage student interaction
Does a reply really answer the question that was posted?	Do I need to answer the question again or clarify further Are students asking good questions and/or helping each other?
What are student misconceptions?	Tailor my intervention in the flipped lab What information needs to be added to lectures
Where do students have confusion [while watching a video]	Tells you what you are going to teach in class based on where students are confused Where I should clarify more for next iteration
How much time do students spend on the material?	Commitment and effort in independent learning Tells me how valuable mat'l is to them. Attitude of students throughout the course Did the material do what I thought it would do (in terms of commitment)
Are students revisiting or reflecting on the material at some point?	Know what information is valuable to them Indicators of higher level thinking and deeper reflection Gauge critical thinking
Are students watching videos with enough time to reflect and integrate before class?	Use to talk to students about their study habits

Table 3.5: Instructor questions and why it was important to them.

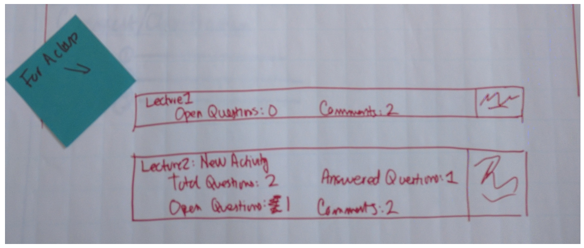
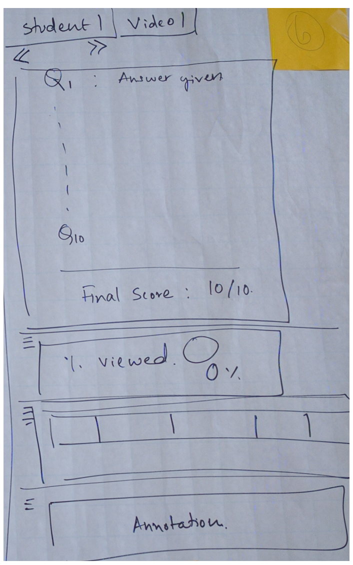
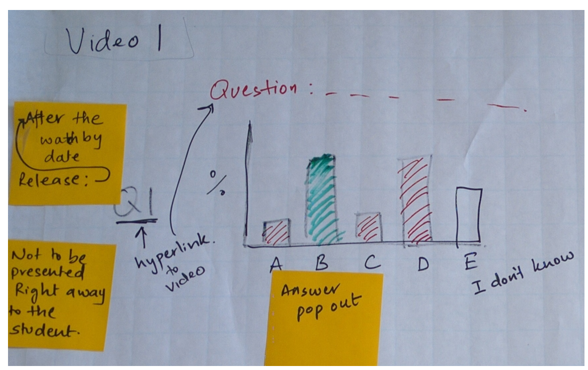


FIGURE 3.4: An analytic to answer "What would get the students more engaged to the content and community?"

them more engaged. The sketch was a textual representation of students' status on the video list page. Students could get an idea for the number of comments they posted, number of questions left to answer, and class averages in comments and questions. Although useful and important to an instructor, there was a lack of design for what the instructor could observe about the student.



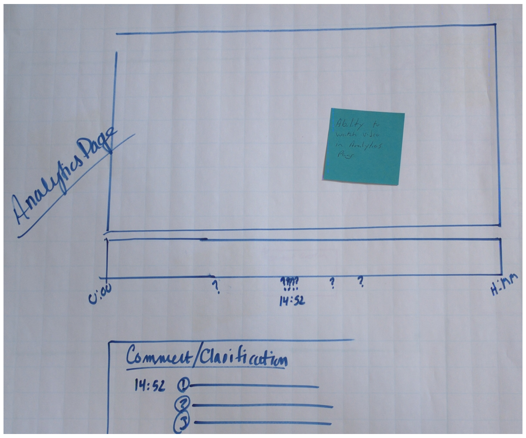
(A) A reorganization of the analytics focused on a single student overview



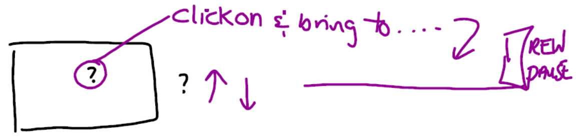
(B) Report quiz results from students in aggregate

FIGURE 3.5: An analytic to answer "What are student misconceptions?"

Figure 3.5 represents the need for instructors to have varying levels of detail in the analytics, and that the current analytics available in TrACE attempt to address this, but do not fully do so. In answering "What are student misconceptions", the rationale for an



(A) "Is this lecture useful and engaging?"



(B) "Where do students have confusion?"

FIGURE 3.6: Two analytics seeking to identify student confusion

aggregate quiz was that the instructor wanted to be able to get a quick overview of student understanding for the whole class. This bar graph closely resembles what would appear in a clicker-style quiz where student answers remain anonymous.

When asking the instructor about what it would be like to view an individual student's quiz, the instructor decided to design a single page that aggregates all of the analytics for one student. Instructors frequently reported that the way they observed students for participation grades was on a single student basis. Often, they have to switch between analytics and remember information in order to have a comprehensive understanding of student activities. Putting all of the analytics in one place allows for the instructor to have a comprehensive view of a single student without being lost among data presented about other students in the class as well. In presenting this information, all of the instructors expressed that allowing for a single comprehensive view would be useful to them.

Although both sessions were conducted at different institutions with instructors that did not overlap in the subject areas they taught, both groups developed a mockup to address student confusion in some way. One group tried to answer the question The question "Is this lecture useful and engaging?" (Figure 3.6a). While the other group was more direct in asking "Where do students have confusion [while watching videos]?"(Figure 3.6b). Instructors felt the need for student to notify them whether they were confused in any part of the video and to provide feedback in a quick and unobtrusive way. Even though students could post questions within TrACE, instructors were concerned that students were not posting questions whenever they had them and wanted a way to provide input with a lower barrier for entry. Both groups had independent ways that they wanted to better understand confusion. One group wanted to know where they can improve on the video content if they notice a cluster of confusion points in a certain part of a video. This manifests itself as question marks along a timeline that may cluster together if many students are confused. The other instructor group wanted to be able to click on a confusion marker and get detailed student activity data (like linking to the session summaries page, Figure 3.2b). The reason for this was that instructors wanted to be able to view what actions students were taking to better understand the material. This information could be used to scaffold students in order to resolve their issues. This follows the previously observed format of having high-level aggregate data but with the ability for more in-depth data if the individual instructor wishes to delve deeper.

3.3.4 Affinity Diagram Results

Work activity notes, which are the elements of the affinity diagram, were created from transcripts, videos, and researcher notes for each participatory design session. A work activity note summarizes a description or complete idea that emerged from the participatory design session. Overall, 84 work activity notes were created from paraphrased mentions of instructor goals, expectations, and use of the system, and features in TrACE. The order of the notes was randomized to reduce bias and organized into categories. The process took a bottom-up approach and grouped similar notes together, only naming the resulting categories after all of the notes had been clustered. A work activity note (phrased from the instructor's perspective) is included with each category. The first major category of notes formed around instructor actions. The subcategories here all dealt with how instructors were using the information they had gathered from TrACE.

Reviewing TrACE data before class- Instructors often refer to the analytics or read student posts in order to gauge what material or questions should be addressed-in class. Instructors usually read through student questions or responses to instructor-crated prompts to determine where extra time should be spent reviewing.

> *"I scan through the video before class to see what students have questions on so I can address those issues in-class"*

Sharing analytics with the class- Some instructors take either an analytic page or an aggregate of various pieces of data and shows it to the class directly. This was done to make students more self-aware of their behaviors as well as reinforce that participating in the system was beneficial to their learning (if correlating behaviors to exam scores)

> *"I share the Percentage Viewed page with students and explain what it means to me in order to give students responsibility for their own learning."*

Other sharing- Instructor actions that did not fit into either group included discussing analytic and post results with Teaching Assistants prior to class so they could more effectively assist students with common questions or printing out data from TrACE (like posts) to give to students as notes in order to encourage more posting.

> *"I use post results to talk to student teachers about what they should focus on during class."*

Instructors also commented a lot on their existing pedagogy such as course goals and student activity within TrACE.

Student Collaboration- This category contains activity notes related to how students should interact with each other and the quality of posts that they create. Instructors often commented on having a "wait and see" approach by not immediately responding to every question so that students had the opportunity to interact with one another. In general, they desired quality posts from students which could help them better prepare for the class.

> *"I want students to interact with each other in TrACE instead of only asking me direct questions and features that would help motivate students to interact with each other."*

Student reminders and self-awareness- How instructors are currently reminding students of what to do as well as their desire to make students aware of what they are doing as well as be more aware of what needs to be done. Instructors discussed sending email reminders to students but wanting student-focused analytics that would help them be more aware of their own actions in the system.

> *"Even good students are forgetful of completing assignments and I waould like a better way to remind them of what to do close to deadlines"*

Goals for analytic use- What goals to instructors have for using analytics. This could include wanting to check student participation, understand student behaviors, or conduct action-based research. Instructors discussed analytics in general instead of specifying any one visualization.

"I use the analytics to know who's prepared, who's not prepared, and addressing students when they are not prepared."

Pedagogy behind videos in TrACE- Examples of the ways TrACE videos are formatted or used (style, frequency of using videos, where they are watched). Some instructors used lecture-style videos to be watched outside of class while others had case studies that were viewed in class. Some instructors had videos for every lecture, while others only used them during brief periods during the semester (such as for lab activities).

"I use TrACE to demo videos to students so we can engage in problem solving together."

There were several categories related to instructor's expectations of students. This could include explicit expectations instructors had for their students (or having none at all) and the challenges that instructors had with ensuring that students were actually meeting the expectations that instructors had set.

Watching Expectations- Expectations related to viewing a video or watching a certain % of the video. Also includes notes that mention analytics specific to viewership. Measuring the percentage of a video viewed (Percentage Viewed visualization) was mentioned the most, with the Viewing Summary analytic being the next most referenced.

"I use Percentage Viewed to see if they newed all of the videos and work backwards to determine the final score"

Posting Expectations- The other group of expectations were related to posting behaviors in TrACE. This could include responding to specific types of annotations like instructor-made comprehension checks or other student questions, asking questions in the video, or using analytics related to counting student posts.

> *"I expect my students to fully watch videos and answer my comprehension checks by the next week."*

Changes to Expectations- How analytic use has changed or how expectations of students have changed either within a course or through multiple iterations of the same course.

> *"The biggest change I made was moving from only having video watching to requiring comments as a way to formatively assess my class."*

When expectations are not met- Some of the challenges that instructors have either in feeling confident whether or not analytics are being met or quality issues. Some instructors report not being sure if the analytics are creating false-positives for student behaviors especially for watching behaviors. Instructors also had problems with the quality of student posts. For both posting and watching, instructors struggle with getting students to meet expectations but still have meaningful interactions and learning gains from the course content.

> *"I'm not 100% sure that students are doing what I want or actually watching the content."*

No Expectations- Instructors sometimes said that they did not have any explicit expectations at all. In these cases, the analytics were just used as a quick gauge of participation.

> *"I do not use TrACE to alter my course content. I just want a gauge of participation."*

Instructors also reported a number of challenges with using the analytics in TrACE. Particularly, instructors reported a lack of actionable output. They often could not figure out how to aggregate the data from the different analytics to create actionable data in the first place or if they could, it was time consuming and took a great deal of effort. Work activity notes that went into this challenge category included *"When information is available to me, I find it hard to aggregate to figure out in general what I should do."*

Request for features that reorganize or display analytics in a meaningful way were proposed by instructors to counter some of those challenges. This feature request category relates very closely my development goals, so each of the notes are written here in their entirety:

- *I want to be able to isolate or specify videos so I can do more, different kinds of analyses.*
- *I want to be able to group videos in the analytics (by section, lecture, or exam) to see why students may have low performance over a time period.*
- *My current workflow is to jump between different analytics and adding date filters. I need a way to integrate these views.*
- *A feature like "View this much(%) by this day" and a checkmark would be useful to me.*
- *I want the analytics to be able to combine info and give me more complex views so I have less manual calculation.*

Some other features were also suggested, but they were unrelated to making the analytics easier to understand. This category covered features like posting or adding quiz questions to videos. Many of these other features are those drawn by instructors in the participatory design session.

3.4 Discussion and Design Requirements

The findings from this section are summarized below in the form of design requirements. I continue this section with the chain of reasoning behind these design requirements. In the next chapter, I detail how these requirements were applied to development of the prototype.

1. Viewing videos and posting are important to instructors. Instructors also use analytics related to percentage viewed and post counts the most, so aggregated analytics should at a minimum be able to aggregate metrics related to both.

   Several instructors did have a preference for some analytics over others. One hypothesis was that some instructors are using analytics in search of specific information instead of exploring the data. The watching and posting expectation categories support this, as some instructor looked specifically to watching or posting analytics that tied to their expectations.

2. Instructors want additional ways for students to interact with videos, so analytics should be able to accommodate those new interaction methods. (As an example, a quiz feature was added to the system, so metrics related to quiz responses should be accounted for)

   In double-looped learning (Argyris and Schön, 1978), initial goals and overall practice evolve as reflection occurs. This not only affects how instructors interact with students, but the design of systems as well. Many design sketches proposed new ways for students to interact with TrACE. Instructors have new questions they want to ask which may require new interventions. Iterating through the learning analytics cycle (Clow, 2012), a changing interventions would alter what data about learners is collected and thus how analytics present that new information.

3. Some instructors rely heavily on video due dates, so information should be filterable by those deadlines.

   Reflective practice involves gradual discovery which informs and changes practice. Without the presence of due dates, there may be no student data present when it comes time for the instructor to look to metrics to inform the next class period. Due dates also allow instructors to have enough time to review questions students may have posted or make other insights necessary for reflection.

4. To ensure accuracy, data provided to students and instructors should be synchronized and provide the same meaning.

   This emerged in the *Student reminders and self-awareness* category, but this also is supported by Nicol and Macfarlane-Dick (2006). In particular, allowing students to understand their behaviors (and correct them), encouraging teacher and peer dialogue, and facilitating self-assessment are some of the reasons why similar analytics should be available to students and instructors.

5. To support a wide range of time instructors can afford to spend with the analytics, they should be effective at a glance with the ability to delve deeper.

   I observed that instructors are most likely limited in the time they can spend with analytics due to the fact that analytic use was done in short bursts that may cover multiple videos. Often, the analytics with the richest data were used for the shortest amount of time or not at all (Session Details). Instructors were given access to the data to perform in-depth inquiry by utilizing the analytics, but this was not usually done. Instructors were split between some spending more time in some analytics over others while others have no analytics that they significantly use over others.

The process of taking an initial glance followed by a deeper exploration can also be supported by reflective practice. A key part of reflective practice is surprise (expectation failure). The impetus for surprise happens when the instructor has existing pedagogical expectations, but an observation is made that challenges those expectations. A surprising observation gets the attention of a practitioner, which then leads to further inquiry. An analytic that allows for an instructor to take notice of unexpected student behaviors but also supports further exploration would, in theory, allow for reflective practice.

6. We cannot assume that an instructor will use analytics in the same way each semester, and frequently changing factors such as available courses, student needs, and course content could affect system use. Analytics should be flexible for these different contexts.

   Usage between instructors varies widely depending on their class load and personal analytic preferences. Between semesters, some instructors are not completely consistent in their use of the system. Frequency of analytic use decreases over time for many instructors, which was especially true from Fall '15 to Spring '16. This could be due to the fact that instructors may not be using videos-based media to inform their teaching as much, or they are pressed for time, with less time to dedicate to using the analytics. This dropoff in use also seems to be true within a semester, as there were work activity notes by several instructors who have mentioned that they do not utilize videos at the end of the course and that the end of the semester was busy for them. Instructors may need both real-time data about their students in busy parts of the semester and easily digestible summaries for when the gaps between sessions increases.

7. To support a variety of information needs, analytics should aggregate data from different metrics/traces based on those information needs.

8. To reduce manual calculation from instructors, analytics should be able to combine data into more complex views

These two design requirements are informed by existing design considerations for learning analytics(Scheffel et al., 2014), and empirical evidence within this study which both point to making analytics quickly accessible. Although the opportunities exist for instructors to do in-depth or exploratory analysis of student behaviors, often times instructors are not taking this time.

> *"And then I was using it initially – just the very initial analytics to say to myself, okay, so who's viewing what? I wasn't looking at the amount of time they were viewing it. It was the Xs and the green checkmarks. That was my focus."*

> *"There are a little bit complex metrics and sometimes I don't have the time to process what you're telling me. So I tend to use them very infrequently. Number of annotations*[Annotation Summary] *and Percentage Viewed is what my grading is based on. So I do tend to get a quick glance at that in order to make an annotation or not and if they each watched the video or not. So it's very nice, easy metrics."*

As observed with TrACE, instructors almost always use an analytic for under two minutes with a median of around 30 seconds. Even so, this was not true for all instructors, as it can be shown that even though all of these instructors volunteered to use TrACE, system use varied widely enough that designs should address both ends of the spectrum. Many of the analytic drawings that emerged from the participatory design session were organized in such a way that the instructor could glance at and gain information about the course quickly, or could be customized to their needs.

> *"If the viewing summary was like a combination between things like Percentage Viewed and Recency, this would be the only thing that I really would need. If I could set a date, like have they watched it by this date, or at least watched part of it by this date, which Recency gives, and I could also set a percentage threshold. Like 95% if they watch, 95% of it, part of it has been by this date then give them the checkmark."*

9. Instructors often consider their course in units or by milestones and many may be viewing multiple videos per analytic visit, so analytics should be able to analyze videos as groups instead of one at a time.

This design requirement came from both the qualitative and quantitative portions of this study. Two of the questions posed by instructors (Are my students confident in the subject matter? and Are students revisiting or reflecting on the material at some point?) resulted in designs that revolved around important milestones or units in the course. From the *Sharing Analytics* category, one instructor reported reanalyzing groups of videos related to an exam and manually aggregating data from those videos to see if there were any trends that had an impact on student academic performance.

Chapter 4

Analytic Development

The next step for this study was to utilize the information gathered from previous versions of TrACE and the results of the formative study to inform the design of a new analytic that allowed for both students and instructors to better understand the expectations for the course. This prototype was developed with two goals in mind: *(i)* a dashboard that affords and supports reflective practice and *(ii)* visibility of expectations for student use.

Creating analytics centered around an instructor's pedagogical expectations set for students in the class supports many of the design requirements presented in the previous chapter. First, expectations are unique to each instructor. If analytics are designed around unique expectations, these analytics by extension would also be customizable. Results of the qualitative study showed that expectations can be complex and related to multiple behaviors (i.e. watching and posting). Presenting analytics based on expectations aggregates data from multiple sources. Instead of looking at Percentage Viewed followed by Annotation Activity and Recency, a single analytic could combine watching expectations, posting expectations, and deadlines into one location. Analytics should also support a high level overview as well as in-depth analysis. Expectations can be binary; either an expectation is met, or it is not. Knowing whether or not students are meeting expectations serves as the high-level overview, and could bring attention to students who are not meeting an instructor's goals much more easily. Finally, expectation-centered analytics supports student use as well. Transparency and synchronized analytics between instructors and students was a design requirement. Normally, expectations are specified outside of the system such as through a syllabus, reminder notifications, or in-class discussions. Expectations specified by an instructor could be made available to students within the system and would support student self-awareness.

The prototypes were developed using an iterative design cycle common in user-centered design. Taking the initial findings from the previous formative study and quantitative data on use of TrACE, paper prototypes were developed and informally evaluated by local instructors that use TrACE and the new analytic was deployed in Fall 2016.

The expectation analytic had three major components:

1. A way for instructors to specify their requirements and connect them to course content (videos).
2. A way for students to view what was expected of them for the videos that do have specified requirements.
3. An analytic for instructors that presents the extent to which students have met the expectations defined in #1.

This chapter continues by summarizing which findings from the previous chapter informed the design of each of these components, changes that may have been made as a result of informal evaluations, and examples of the final prototype that was deployed in the field study.

4.1 Specifying Expectations

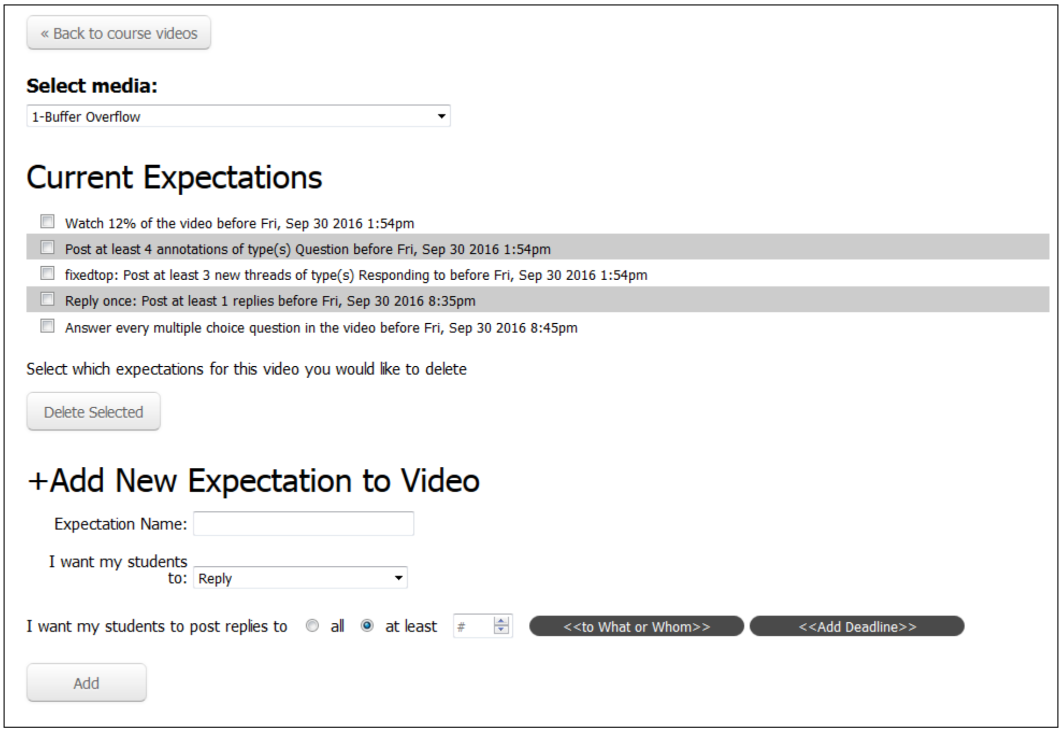
In order to specify expectations, it was necessary to understand what instructor expectations were within TrACE. From the formative study, some of the expectations that emerged include watching the videos in full, watching videos prior to the set due date, posting comments and questions using TrACE's annotation feature, and responding to any instructor-prompted questions scattered throughout a video. Instructors wanted to ensure that students were prepared for class and instructors also wanted to be able to respond to questions that students had. Below was a summary of the design guidelines informed by the formative study in Chapter 3:

- Viewing videos and posting are important to instructors. Instructors also use analytics related to percentage viewed and post counts the most, so aggregated analytics should at a minimum be able to aggregate metrics related to both.
- Instructors want additional ways for students to interact with videos, so analytics should be able to accommodate those new interaction methods. (As an example, a quiz feature was added to the system, so metrics related to quiz responses should be accounted for)
- Some instructors rely heavily on video due dates, so information should be filterable by those deadlines

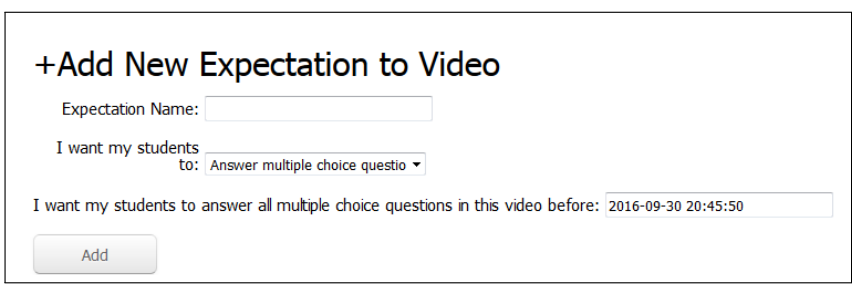
The components of an expectation, from this feedback, include the type of expectation, how much the instructor wants students to do (watch %, post count), and a due date. One design idea was to formulate expectations in the same language that instructors were using to describe them in the focus groups. To evaluate this design idea for specifying requirements, instructors were shown a fill-in-the-blank form detailing expectations for watching, posting, and question-answering behaviors (Appendix A). While the multiple-choice question feature had not been used by instructors yet, all other expectations in this form were ones that instructors would have exposure to in the analytics. Instructors were asked to fill out the form based on their expectations of students in one of their existing (or previously taught) classes. Instructors were allowed write anything they wanted in the blanks, so they were not limited by preset input types. After educators filled out this form, they were asked two informal Likert scale questions about the ease of use in filling out the form and the extent to which this form would cover the expectations they currently have for their course. Instructors were also asked why they gave those scores.

They expressed that the fill-in-the-blank style of presentation for the analytics was easy to understand. In observing how instructors were entering their expectations, many instructors also had “fuzzy” expectations. That was, it was easier for instructors to say “watch between 80 and 100 percent of the video” instead of stating a single hard number. This was also expressed in ways such as “respond to all of these posts” or “watch before the video due date” where the video due date was automatically calculated instead of a manually entered value. Some instructors had difficulties thinking of their expectations in the context of a single video and had a desire to specify expectations for groups of videos instead of a single one. These could be described as macro expectations (tied to the overall course) and micro expectations (a single video). Instructors justified this through wanting to make sure that students did not post low quality content in an attempt to meet the expectation.

Following the fill-in-the-blank format in the prototype, instructors could add expectations from a dropdown menu. Their options reflect the range of expectations mentioned earlier: Watching, Posting (Posting new threads, replying to existing posts, or both), and answering Multiple Choice Questions. The options the instructor could fill in would change depending on their selected expectation type, but the text reflects the type of expectation (Figure 4.1).



(A) Adding Reply Expectations



(B) Adding Multiple Choice Expectations

FIGURE 4.1: Examples of how expectations could be specified

There were several optional elements that instructors could add to an expectation. First, every expectation had the option to have a unique deadline. This could be used in cases such as the video due date being at midnight, but instructors wanting students to fulfill expectations at the start of class or before a big exam. Otherwise, it would default to the video deadline. If the instructor set no deadline for the video, it defaults to the end of the year (a time when any semester would most likely have ended). Both in the formative study and in the informal evaluation, instructors described various types of posting expectations. Thus, instructors could also set what kind of posts they want students to make based on annotation types which were unique to each class. An instructor could ask students to post questions, or have a specific scaffolded annotation type that they want students to use. When replying, instructors could ask students to reply to certain types of posts or users. All of the following were possible expectations an instructor could create with my tool:

- Watch at least 95% of the video before Fri, Oct 21 at 10:00AM
- Post at least 1 annotation of type(s) Comment or Question before Fri, Oct 21 at 10:00AM
- Post a reply to all Comprehension Check and Reflection posts before Fri, Oct 21 at 10:00AM
- Post at least 2 replies in response to other students before Fri, Oct 21 at 10:00AM
- Answer every multiple choice question in the video before Mon, Aug 29 at 12:00AM

I also added features to remove expectations as well as import expectations from other videos. To support changing expectations as a result of reflection, instructors could disable expectations that no longer align with their course goal. Importing expectations was a useful feature for instructors who had limited time to create expectations. However, importing expectations could lead to instructors duplicating expectations without considering how their goals or expectations had changed. This would be evidence of single-loop learning, where practitioners are not engaging in reflective practice. To mitigate this, importing must be done for each video (an extra step) and presented each expectation to the instructor prior

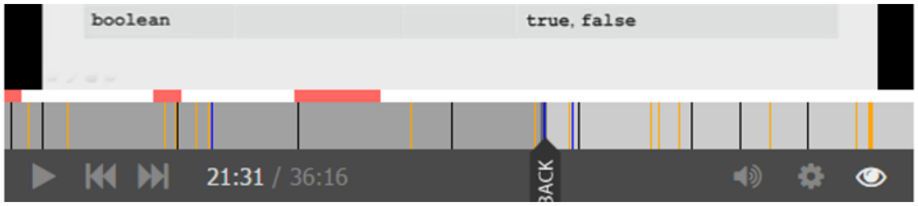


FIGURE 4.2: Feedback to students on their video viewing was visualized above the play-head.

to importing. This made importing expectations less convenient, but encouraged more reflective practice than copying a single expectation to every video in a semester in one step. With expectations specified by instructors, analytics for students that mirror these expectations were then made available.

4.2 Student Analytics

The second goal was to allow for students to see where they were in relation to the class goals set by the instructor, and to have a clearer understanding of expectations. Early iterations of TrACE had few indicators for whether or not a student was meeting expectations such as watching a video at all or how much of the video was already viewed. Through focus groups and feedback, the most recent iteration had implemented a way to show students the percentage of video viewed (Figure 4.2) and an email reminder informing students if they have watched a video before the due date, but this was the only feedback available to students.

To summarize, only one design guideline initially emerged from the formative study that informed these student analytics.

- To ensure accuracy, data provided to students and instructors should be synchronized and provide the same meaning.

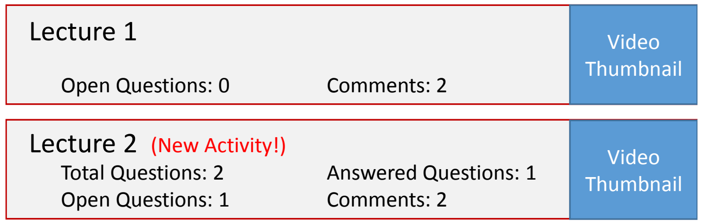
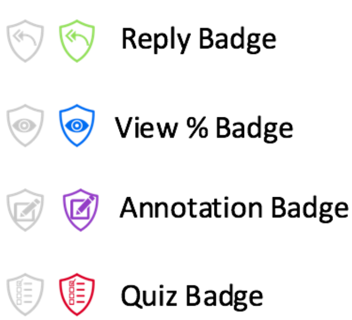
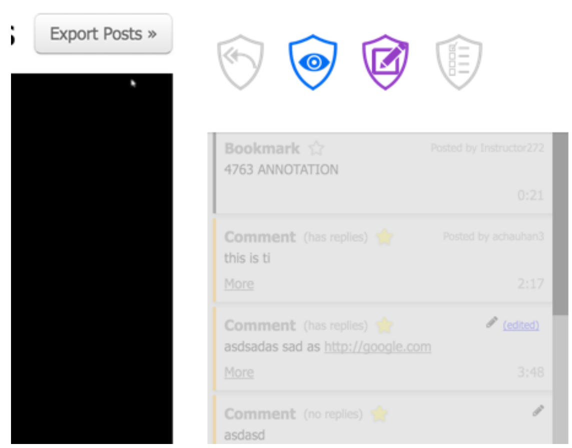


FIGURE 4.3: An analytic designed by instructors to answer "What would get the students more engaged to the content and community?" Redrawn for clarity.

The initial mockups for the student view stemmed from one of the designs in the participatory design session. The design group wanted to engage students with the content and community by showing students their own status on videos. Figure 4.3 is a redrawn version from that initial sketch. This was then re-imagined as visual badge icons to represent expectations. It was also moved from the course page into the video page to allow for students to see their own progress in real-time. Instructors were shown mockups in the form of overlays on the existing TrACE web pages (Figure 4.4).



(A) Initial badge designs



(B) Mockup of badge placement on the video page

FIGURE 4.4: Initial mockups of the student-facing analytics

A pitch was provided to instructors detailing how the expectation analytic would work, how it would be accessed by instructors and how it would be accessed by students. Instructors were free to ask questions about the design during this time. Instructors were asked what aspects of the design were useful, what they liked about the design, and any changes they thought would make it more usable for their course contexts. Once again, the designs were favorable with instructors. Some suggestions included letting students know before entering the videos which ones still had not been completed, so students would know which videos to re/visit in the first place. Also, instructors noted that students might not come back to a video, so letting them know before they leave if they have or have not met all the expectations would be useful. These suggestions became two additional design requirements:

- Students should be reminded of what to do close to deadlines.
- Students do not frequently revisit videos, so the analytic should inform students of their current progress within the video page.

In the prototype, the video list page was modified to show students which videos have met/unmet expectations. This simple view presents a check or exclamation mark based on whether or not the student had met all of the expectations in that video yet (Figure 4.5).

Figure 4.5: Student can see indicators of videos with met or unmet expectations on the course page

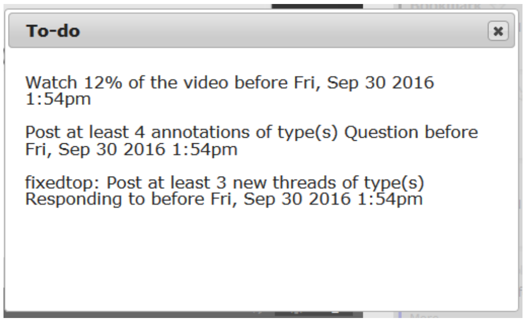


FIGURE 4.6: A To-Do list detailing what the expectations are at the start of every video

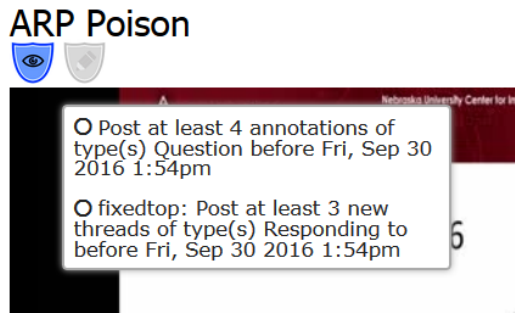


FIGURE 4.7: Students see several badges related to the expectations they should meet

When opening a video, students were presented with a To-Do popup, which lists every expectation the instructor had set for the video in the same language the instructor used to make the expectation (Figure 4.6). A set of badges was located above the video (Figure 4.7). These badges were indicators based on those expectations, and they change from grey and translucent to opaque and colored when all of the expectations in the category (watching, posting, replying, or multiple-choice answering) were complete. Students could hover over each of the badges to see the same expectations that were listed in the initial popup and whether or not they have met that specific expectation. These badges update whenever the student pauses the video, creates a post, or answers a multiple-choice question; students receive live feedback for their actions. Additionally, any changes the instructor makes to expectations were reflected in the badges. If an instructor removed or added more expectations while a student was watching a video, the badges the student could see would update as well, so a student would always have their progress to the current expectations visible to them.

4.3 Instructor Analytics

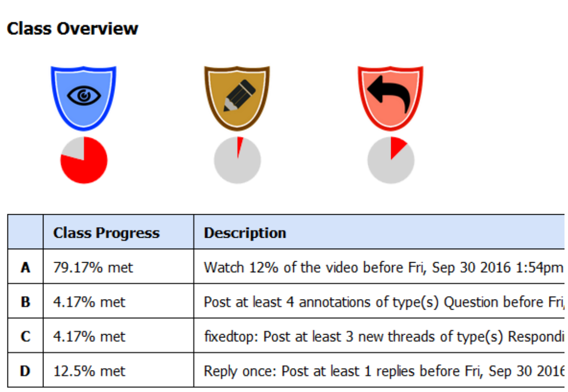
Instructors specified expectations which are then made visible to students. The final part of this cycle returns to instructors by allowing them to see how well students have met their expectations. Before, instructors needed to access different tabs/pages in order to understand student behavior in the system. Instructors expressed they had challenges with workflow consistency because of the frequent context switching. My goal for this portion of the analytic was to aggregate data from multiple pages in one location instead of the previous setup where multiple pages on the dashboard provided different information that must be gathered and cross-referenced by the instructor. Results from the formative study also informed this portion of the analytics:

- To support a wide range of time instructors could afford to spend with the analytics, they should be effective at a glance with the ability to delve deeper.
- We cannot assume that an instructor would use analytics in the same way each semester, and frequently changing factors such as available courses, student needs, and course content could affect system use. Analytics should be flexible for these different contexts.
- To support a variety of information needs, analytics should aggregate data from different metrics/traces based on those information needs.
- To reduce manual calculation from instructors, analytics should be able to combine data into more complex views
- Instructors often consider their course in units or by milestones and many may be viewing multiple videos per analytic visit, so analytics should be able to analyze videos as groups instead of one at a time.

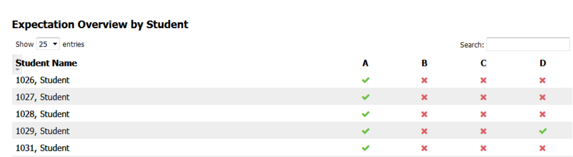
Feedback was gathered from instructors on paper prototypes of this portion of the analytics at the same time as the student-based analytics. First, when looking at a single student in the analytics, instructors requested that we make it easier to cycle between students and videos (possibly with previous/next buttons). Also, instructors thought it was important to give clear feedback about whether or not an expectation had been met in this same detailed report. This way, instructors instructors would not be burdened with mental calculations to determine whether a not a student met an expectation.

In the prototype, the instructor expectation analytics were divided into two new analytics to support instructors who desire that high-level overview with the option for more detailed information. The first, the Expectation Progress Report (Figure 4.8), represents the high-level information an instructor would like to know about students and the class overall. The Student Report stems directly from the participatory design session and offers a way for instructors to combine all of the analytics on a single student in one page. Both of these analytics were described in more detail below.

4.3.1 Expectation Progress Report



(A) How did my entire class meet my expectations?



(B) Did individual students meet my expectations?

FIGURE 4.8: The Expectation Progress Report

When the instructor selected a video within their course, they first saw a high level overview of how all students in aggregate were doing regarding watching, replying, posting, and multiple choice expectation categories (Figure 4.8a). Instructors first saw a pie chart for each category (80% of my class met my watching expectations) and below could see a breakdown of how well the whole class did for each individual expectation. Instructors were reminded what their expectations were by using the same language they used to define those expectations in the first place. From there, instructors could scroll down to see an overview of each student in their course (Figure 4.8b). This was similar to the Viewing Summary report, by also using checks and X's, but instead of being based on a metric that instructors cannot modify, the checks and X's relate directly to what the instructor had specified as their expectation.

A link was available here to the Student Report, where instructors could get even more detailed information about individual students, if this overview was not detailed enough for the instructor.

4.3.2 Student Report

The Student Report analytic was developed separately from this study as a direct influence from one of the designs that emerged from the participatory design session. This arose as instructors expressed a desire to see student details all in one place without jumping between analytics. The Student Report had several features. First, it contains analytic panels, which could be minmized and reorganized based on what the instructor prioritizes in viewing analytics. Each panel was a student-centric version of the existing analytics. For example, an instructor could see the percentage viewed, media activity, or post counts for a single student. Additional panels available to instructors include seeing a student's multiple

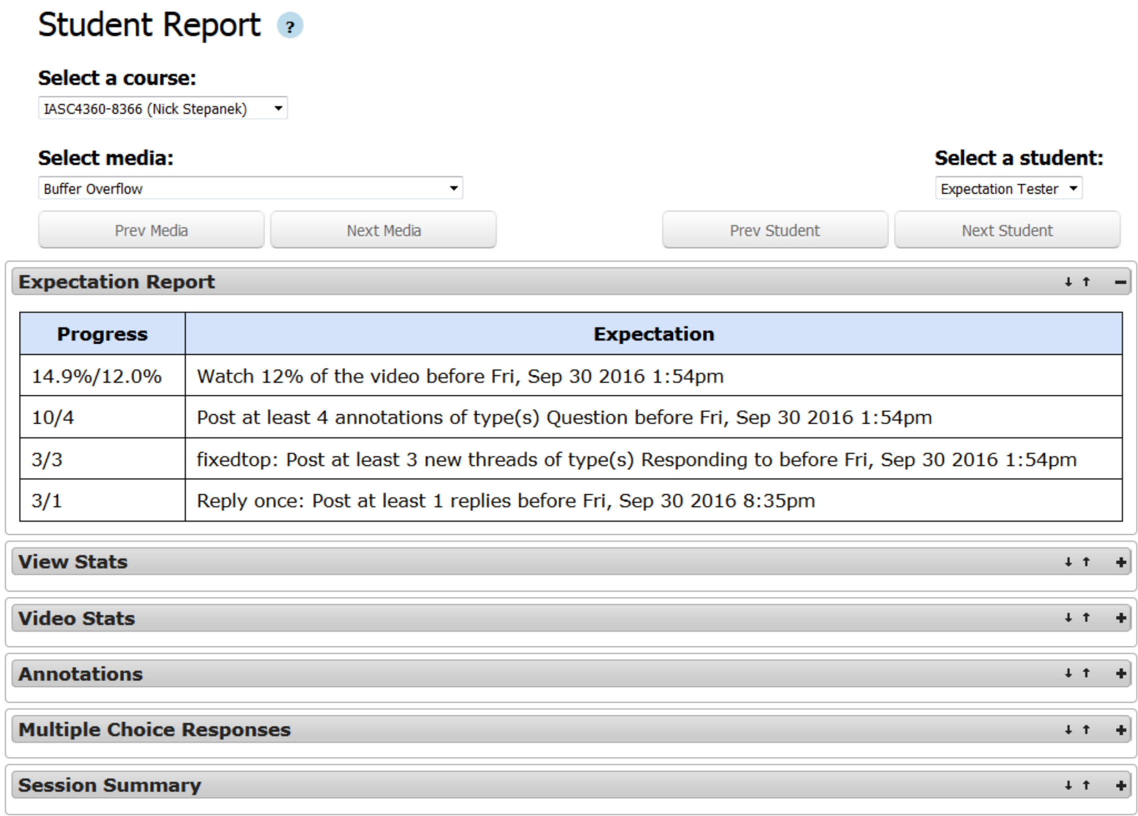


FIGURE 4.9: The Student Report Analytic with the Expectation Report panel open

choice responses, posts, and reply content for a single video. This allows for the instructor to get quick insights into the quality of a student's posting activity without having to open the video and manually search for that student's contributions.

My contribution to this analytic was the Expectation Report panel (Figure 4.9), which was displayed at the top of the Student Report. For every expectation the instructor had, this panel informs instructors to what extent the student met the expectation. From this, instructors would know how close the student was to meeting the expectation, or how far beyond expectations the student performed. If this panel sparks any inquiries, the instructor could easily view another analytic panel for more detailed information.

Chapter 5

Field Study

With the implementation of the prototype developed from the previous iteration, the final phase of the study was to evaluate it with both students and instructors. Although the majority of learning analytics evaluation work focused on improving design of tools (Dyckhoff et al., 2013), the goal of this field study instead looked at evaluating the behavioral impact on teachers and students. The remaining research questions were answered by this evaluation.

- **RQ3**-How do instructors instructor behaviors change with the explicit presence of this analytic?
- **RQ4**-How do student behaviors change with the explicit presence of this analytic?

This field study was conducted during the Fall 2016 semester and included all 10 instructors teaching 20 courses that use TrACE. 70% of classes have video content imported from previous iterations of the course.

To measure the potential impact of expectation-centered analytics, a within-subjects study design was applied to both instructors and students. The rationale behind using a within-subjects approach was that each instructor taught a course (or multiple courses) with unique contexts that does not afford comparisons between instructors. Additionally, the population of instructors using TrACE was fairly small, so a between-subjects comparison would lack statistical power. Although there were a large number of student users, the individual class sizes were much more modest (around 30 students per class, not considering consent rates) so a between-subjects approach would be insufficient in this context. By using a within-subjects approach, individual differences in students' overall levels of performance would be controlled. Also, students may not complete the course, so comparison of a student to their own behaviors ensures that there would not be complications if other students drop the course. A possible confound in analysis was that technical changes to the system were not limited to the addition this new analytic. Other analytics have been developed and were deployed in parallel to the expectation-centered analytic. It may be difficult to attribute improvements explicitly to one analytic, but the use of a within-subjects approach should control for this effect, as between periods only the expectation analytic was introduced. A limitation to using a within-subject methodology was the possibility of carryover effects which could bias the treatment period, but informing instructors and students of system changes could mitigate this.

The first phase took place from the first day of class and lasted approximately 4 weeks (this varies slightly by course start date) The second phase was also 4 weeks. During the entirety of the study, instructors were asked to use the system to log their expectations of students. In the control, neither instructors nor students received feedback through expectation-centered analytics, and the courses ran the same way as in previous semesters. In the treatment period, the analytics were enabled and introduced to all users. Table 5.1 was an overview of the format of the field study.

Table 5.1: Overview of the phases in the field study

Group	**Control (Start-9/19)**	**Treatment (9/19-10/21)**	**Analysis**
Instructors	Specify Expectations for videos		- Details on expectations
	- Student Report (No Expectation Panel) - No Expectation Report	- Expectation Panel in Student Report - Expectation Report	- Within-subjects comparison of analytic use
Students	- No notifications of expectations - No badges in video - No To-Do List	- Expectation Notifications - Badges visible in video - To-Do List in video	- Within-subjects comparison of performance

The number and type of expectations created and the changes/updates to these expectations that occurred over the course of the study were logged through TrACE and analyzed. Data collected on creating expectations included the type of expectation and details, when the expectation was made, if it was imported from another video, and if it had been disabled/deleted by the instructor. From this information, we could determine what kind of trends instructors have. For courses that already have videos imported from other semesters, do instructors set expectations once for an entire course and leave expectations alone for a semester? Or would instructors periodically revisit their course expectations and modify them as the semester progresses?

Overall, 7 of 10 instructors had created expectations during both phases of the study and 238 expectations in total. Of these 7 instructors, the majority of expectations were posting expectations (59%), followed by watching expectations (31%) and finally quiz-answering expectations (10%). 76% of these expectations were imported. Importing expectations is a behavior that does not show evidence of reflective practice. When importing large amounts of expectations, instructors may not be considering how those their practice and thus their expectations, change over the course of the semester. In general, instructors had between one and two expectations per video ($\mu = 1.36, std = 0.52$), and for the most part these expectations remained unchanged throughout the study period. This indicates that if reflective practice took place, it did not manifest itself through evolving expectations. Instructors would either do mass-uploads of expectations at the beginning of the semester, or they would gradually add expectations as each video was released to students. Although instructors had the ability to change the deadline of expectations, most chose to use the default, which was the same as the video due date.

5.1 Instructor Analytic Use

The evaluation of RQ3 was done in two ways. First, instructor activity data in the analytics was collected from the 9 instructors who had set any expectations at all throughout study. The method of collection was the same as the methods detailed in chapter 3. Summarized again, click-level actions were logged when an analytic was opened in TrACE. Data collected included which analytic was accessed, the course, timestamps, and other action details (such as applying filters). The actions generated by instructors were grouped into *"sessions"* that were cut off after 15 minutes of inactivity. Sessions shorter than one second were filtered out. Within a session, an instructor could look at multiple courses (if teaching more than one), videos, and students.

To understand the extent to which instructors used the tools provided and in what ways these behaviors have changed, two main metrics were used. **Frequency** of analytic access and **duration** of analytic use. Duration was not normally distributed, so a Wilcoxon signed-rank test was used to compare the differences in time spent overall between phases, and time

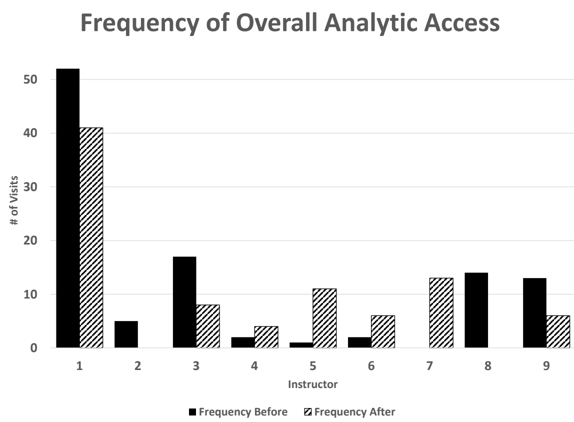
Table 5.2: A contingency table comparing analytic use (frequency) for instructors 1, 3, and 9 between phases

Instructor	Phase	Annotation Summary	Media Activity	Percentage Viewed	Quiz Analytic	Student Report	Loyalty	Recency	Expectation Report	Viewing Summary	View Count Graph
1*	1	15	1	13	1	11	1	1	0	8	1
	2	8	1	9	0	2	0	0	17	3	1
3	1	2	2	1	6	3	0	1	0	1	1
	2	0	0	0	1	4	0	0	2	0	1
9	1	1	0	1	2	8	0	1	0	0	0
	2	0	0	2	0	3	0	0	1	0	0

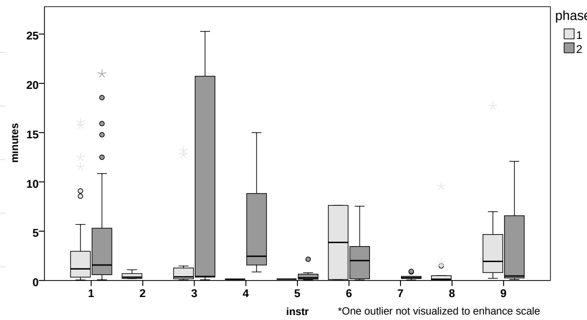
spent with each analytic to see if there were any differences. For frequency comparisons, a Fisher's exact-test was used, as the expected frequency for any given analytic in the contingency table was expected to be less than five.

Figure 5.1a shows the frequency of visits before and after for instructors. Using a Fisher's exact test comparing the ratio of analytic between phases for a given instructor, only one instructor had a significant difference ($p < .001$) in the analytics accessed (Table 5.2). When the treatment period began, 41% of Instructor 1's analytic visits were in the new expectation report. Figure 5.1b shows the differences in duration. Wilcoxon signed rank tests did not show any significant differences in duration ($Z = -.652, p > .05$) between the phases for instructors. Only three instructors used analytics at least 5 times in each phase, so the small sample of visits could be a reason for a lack of statistical significance. The median time that instructors spent with the analytics was still not incredibly high, with most instructors still spending less than a minute in the analytics. However, there were cases, such as when using the student report, where instructors who did use this report spent a significant amount of time looking into each student's behavior.

The presence of this analytic did not encourage all instructors to spend more time with analytics in a significantly different way. Although, one instructor, who was already a heavy user of analytics, quickly adopted the use of the Expectation Report analytic when it was introduced.



(A) Frequency of visits, before and after



(B) Duration (in minutes) before and after. One outlier for Instructor 9 excluded to enhance scale

5.2 Student Performance

To evaluate RQ4, a quasi-experimental design was applied. Students were compared between the control and treatment periods using a Wilcoxon signed rank test for system activity. Performance was defined here as the extent to which students met an instructor's expectation for a video. This was simply calculated as (student completion amount / required amount) Student completion was calculated from the amount that a student watched, posted, or answered quiz questions before the deadline specified in the expectation. The required amount was provided by the instructor when creating the expectation. Click-level log data built into TrACE allowed for these student behaviors to be calculated.

Courses of instructors who did not specify expectations in both phases were excluded, as well as students who were reported as having dropped the course. 188 consenting students among nine courses remained. Average compliance to all expectations in each period was calculated and a Wilcoxon signed rank test was used as a pairwise comparison of student performance. Another question that rises from this analysis was "do students meet some types of expectations better than others?" To answer this, the extent to which students in a course met question, watching, posting, and quiz-answering expectations was compared

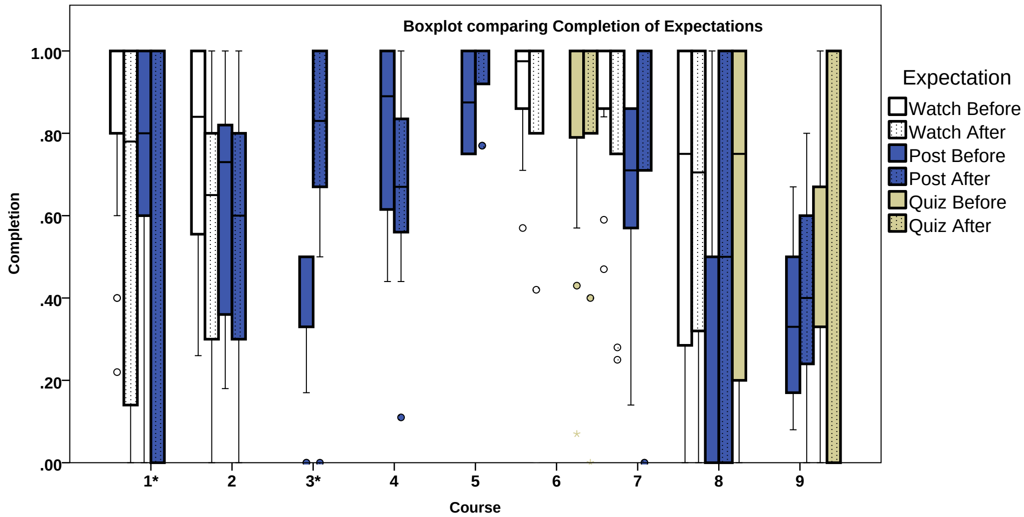


FIGURE 5.2: Average completion of expectations (up to 100%) compared between phases separated by expectation type. Classes with significant differences in performance between the control and treatment($p < 0.05$) are marked by *.

to other expectation types in each period.

Only two courses had a statistically significant difference between one phase and the other (Figure 5.2). Course 1 met significantly fewer expectations from the control to the treatment ($Z = -2.38, p < 0.05$) going from a median of 85% to 48.4%, and Course 3 had a significant increase in performance from the control to the treatment ($Z = 3.75, p < 0.01$). Looking more closely at Course 1, it was the watching expectations that had significantly lower completion ($Z = -2.8, p < 0.01$), with no difference in posting expectations ($Z = -0.284, p = 0.78$). The latter half of the semester was student-created content which was not as strongly incorporated with the class as with instructor-created content. The large number of videos required to be watched at the same time could have lowered the completion rate of these expectations. With Course 3, there was a significant increase in completion of posting requirements, rising from a median 50% of expectations met to over 83% met.

With only one course seeing any improvement, I can conclude that students did not complete more expectations overall with the introduction of expectation-centered analytics.

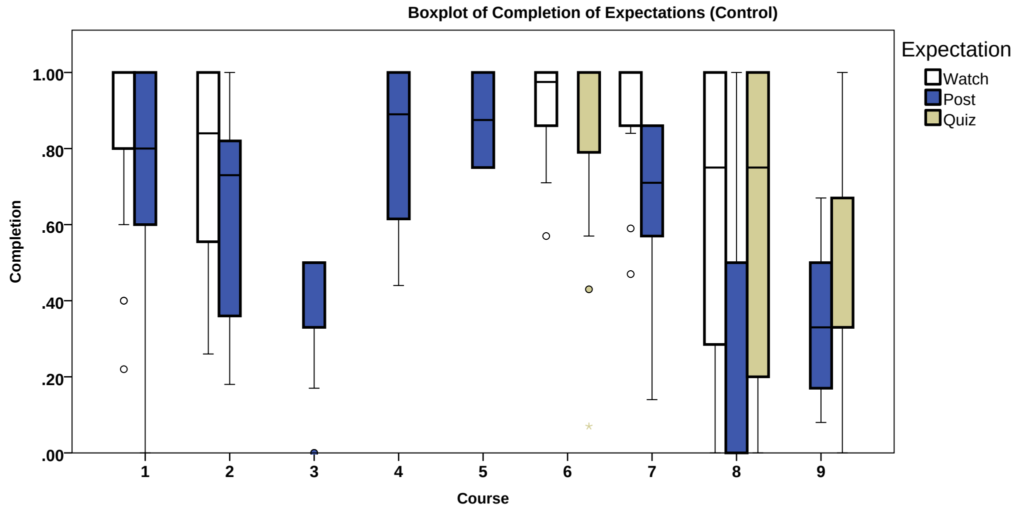


FIGURE 5.3: Average completion of expectations (up to 100%) in the first phase, separated by expectation type.

However, in observing differences between expectation types, I did start to see an interesting trend (Figure 5.3). From the Wilcoxon sign-rank results comparing expectations types to each other, when expectations were not visible to students (control), the posting expectation was *always* significantly lower than other expectations ($p < 0.05$) in any given course. There was no difference between completion of watching and quiz expectations for any classes. It could be that of the three, posting was the most difficult expectation to meet for students, or students would watch a video without actively collaborating, even if that was the expectation of the instructor. Quiz questions paused the video as the student was watching, so if a student was faithfully watching all of the video, they would encounter all of the quiz questions along the way. The same could not be said for posting, where a student may need to seek to a point in a video or find a post to reply to.

In the second phase, there was no longer a significant difference between posting and watching/quizzes (Figure 5.4) except for Course 9, where students still were worse at posting (median 40%) than answering quiz questions (median 100%) ($Z = -1.9, p = .047$).

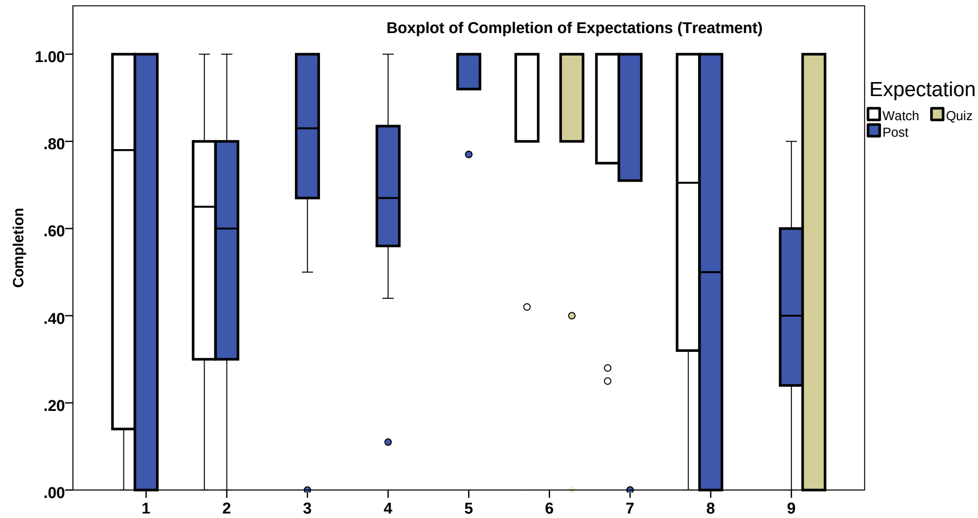


FIGURE 5.4: Average completion of expectations (up to 100%) in the second phase, separated by expectation type.

Although there was not a significant difference, there was an upward trend in student posting behaviors for the majority (6) of classes in the treatment period. There may be an improvement here, but there may not be enough participants to make it statistically significant. Future studies on courses with larger class sizes may provide more insight on whether or not expectation analytics have a positive impact on student posting behaviors.

Chapter 6

Conclusion

In this thesis, I conducted a three-phase study consisting of *(i)* a formative study of instructor analytic use in TrACE, *(ii)* development of an expectation-centered analytic, and *(iii)* a field study on the impact of this expectation-centered analytic on instructor and student behaviors. My hypothesis was that a learning analytic that encodes and reifies instructors' individual expectations would better support reflective practice for instructors and allow students to more reliably meet set expectations. The research questions that motivated and informed development and evaluation of this learning analytic were:

- **RQ1**:How do instructors currently conduct inquiry on student behaviors?
- **RQ2**:What expectations do instructors see as valuable to model within the context of learning analytics?
- **RQ3**:How does instructor inquiry change with the presence of this analytic?
- **RQ4**:How do student behaviors change with the explicit presence of this analytic?

To summarize the findings of the formative study, results showed that instructors had very different behaviors, needs and expectations. Analytic use in general occurred in brief sessions less than a minute long. Some instructors prioritized using analytics related to their goals, however some goals were beyond quantitative measurement. Some instructors looked to the quality of student understanding, and in this case numerical analytics would not be useful. When asked about these behaviors, instructors reported not having time for in-depth analysis. Also, instructors reported that although there was data available in the analytics, they did not know how to make that data actionable. Instructors often thought of their expectations and inquiry by course unit, as opposed to the current organization of TrACE which is by video.

Instructors had expressed a desire to model their expectations and to allow for students to see analytics. These needs motivated the development of the expectation-centered learning analytic. The learning analytic in TrACE was built as multiple parts. Instructors specified expectations, students could see those expectations both on the course page and within a video, and instructors could see the results of student activity in the Expectation Report and the Student Report analytics.

The results of the field study did not support the hypothesis. An essential part of reflective practice involves gradually exploring data to come to an understanding of the situation. If this process is taking place, it would be expected that instructors would use the analytics to follow-up on surprising observations discovered through the analytic. However, instructors for the most part did not change their behaviors with the introduction of these analytics. One did, but this instructor had been a consistently heavy user of analytics.

It was proposed that considering the Learning Analytic Cycle in development, especially by building metrics with intervention in mind, would support instructor reflection in turn (Clow, 2012). Once again, evidence is not in support of reflective practice taking place. One expected outcome, should the metrics have influenced instructor interventions would be a change in expectations. Expectations arise from instructor goals, which would change with reflective practice. I did not find evidence of changed expectations, so if reflection did take place, it may not be perceptible from expectations alone. High import rates for expectations also fails to support the hypothesis. Imported expectations reflect what the instructor's goals were at the time of initial creation (which may even be from the first video at the start of the term) as opposed to an instructor's current goals. Instructors also did not always set expectations for videos, which means that goals were not explicitly considered in the first place. There may be other variables at play which limit the extent to which instructors can reflect on their practice. Areas for further exploration are elaborated upon in the Future Work section.

Students did not meet expectations more reliably with the explicit presence of these analytics. Only one course saw a significant improvement in performance. It is interesting to note that without explicit expectations, students were significantly worse at meeting posting expectations than anything else. However, with explicit expectations, posting was no longer worse than watching or answering multiple-choice questions

6.1 Limitations

There are several limitations and confounds which could have possibly affected these results. First, small sample sizes for both instructors and students(in a course) limited the statistical power of all analyses. This is a fundamental challenge and trade-off of using a small research-based system such as TrACE. Introducing the system to larger class-sizes could improve support of RQ4, specifically. I tried to mediate the effect of sample size by using a within-subjects methodology. A possible confound introduced here is a natural loss of motivation (and thus not meeting expectations) as the course progresses.

Finally, some changes to study design may have improved the quality of results. First, while the mixed-approach formative study gathered many points of data, a Participatory Design session, especially a remote one, was difficult to execute and interpret. I attempted to compensate by having field notes, a scribe, and video recordings, but because of excessive cross-talk, the focus groups only partially had verbatim transcripts to work from. Even a small change such as moving participants to opposite sides of the rooms during individual work would have greatly improved the quality of transcripts. Second, all coding of data is improved when there are multiple coders. I created the work activity notes and affinity diagram independently, which could have introduced some of my own biases into the resulting themes.

There were some limitations to analytic design, and some features in the analytic did not match up to the design guidelines mentioned in Chapter 4. My analytic was designed so that instructors specified their expectations for each video individually instead of a group of videos. This design choice seems counter to how instructors organized their courses on a per-unit basis instead of on a per-video basis. This was an attempt to encourage instructors to consider and change their expectations more often. Additionally, the underlying analytic systems in TrACE were organized on a per-video basis, so major changes to how the system and its dashboards function were outside the scope of this study. Future iterations of this prototype, should it be useful for instructors, could be modified to allow instructors to organize expectations or analytics by unit instead of by video.

While conducting the field study, there were several system-wide changes in the Fall 2016 semester that were not introduced in previous semesters. First, the Student Report (without the Expectation Report panel) was made available to instructors alongside a new Quiz analytic. Quiz questions was also a new feature introduced, so instructors may not have fully integrated it into their classes. There were also some bugs which could have impacted a student's ability to meet expectations. TrACE is deployed at multiple institutions across two different time zones. Time-related issues briefly caused expectations and quiz deadlines to be shifted one hour earlier. If students in the affected timezone attempted to interact with TrACE within this one hour period, they would not have been able to successfully answer quiz questions. Additionally, expectations do not consider activity beyond the deadline, so the analytics would report to both instructors and students that expectations were unmet. This bug was corrected for expectations, so student activity during that time period was not excluded from analysis. From anecdotal evidence and bug reports, having customizable expectation deadlines did create some challenges for instructors. An instructor could set video, expectation, and quiz deadlines independently from each other, and sometimes these were unintentionally misaligned. Repeating this study in another semester when instructors are more comfortable with features and without system instability could address this limitation.

6.2 Impact and Future Work

Although the hypothesis was not supported, this work does contribute to both education research and practitioners. Although the system was designed with education theory in mind, some limiting factors could have reduced the effectiveness of the system. Primarily, the formative study demonstrated the extent to which instructor time plays a role on analytic use. Instructors are very limited in the time that they spend in analytics, so even minor inconveniences such as context switching, mental calculation, or even unexpected course changes become a huge barrier to analytic use.

This work makes the case for more user-centered practice in learning analytics. This is a look at developing learning analytics not only *for* primary stakeholders (instructors and students) but *with* them using support from both formative evaluations and education theory. This study is a useful thought piece on what incorporating instructors looks like in research and especially for working with experienced learning analytic users instead of new users. Designing for instructors was identified as a need by Dyckhoff et al. (2013) and prioritizing instructor involvement was also a need (Nelson et al., 2008). This study, especially in the formative evaluation phase, can offer valuable insights to future researchers and developers on how to continue to involve instructors in learning analytic development.

Second, this work evaluates the impact of learning analytics not only for student performance, but for instructor use as well. Understanding the behavioral changes a learning analytic has on instructors is a necessary step in integrating learning analytics within course contexts. Even when incorporating instructors early on in the design process, the developed result may not be successful. While many studies evaluated analytics through usability questionnaires or instructor/student opinions, I was able to confirm that this learning analytic did not have a drastic improvement instead of retaining the implicit assumption that it did.

Third, the designs and features instructors came up with revealed a fundamental difference between how the system organized analytics and how courses were organized. Instructors conceptualized their courses as multi-video units instead of as single videos. This emerged from the formative study and applied to both how instructors wanted analytics organized and how they expressed their expectations in the course. This finding was not one that was expected and was not apparent in the reviewed literature on learning analytic design. Further literature review will be needed, and developers of learning analytics and other education-support systems should take this into account early in design in order to better support instructor inquiry.

Although instructors did not conceptualize their courses on a per-video basis, directing instructors to specify expectations on a per-video basis allows for researchers to understand these expectations at a very fine level of detail. I made this design choice to allow for a fine level of analysis throughout each of the courses in the study. In previous studies on student viewing behaviors, some of the only ways to know if there were expectations that could explain student behaviors was to either read instructor journals (of which there may not be any consistent reporting of expectations) or ask the instructor. This study implemented expectations to be explicitly used in the system and thus we had full knowledge of expectations and were able to collect data on the completion of the expectations. Another benefit of making expectations explicit is that we were able to better understand changing student behavior as it was either related to changes in expectations or with unchanged expectations allowed to find interesting changes in student behavior that warranted more investigation.

For practitioners, those with a focus on collaboration and posting may have more challenges with students meeting those expectations than instructors that only have watching expectations. Instructors that value posting may have to take additional steps to support students in meeting expectations. Also, although practitioners thought it would be useful to have analytics available to students, we did not find that making this information available to students changed how well they met an instructor's expectations. Overall, students may have needed more support than what expectation-centered analytics provided. There are some cases where students greatly improved expectation compliance, so knowing what influenced success in those classes could also be used to support other classes as well.

6.2.1 Future Work

One possible avenue of future work is understanding affect on instructors, and especially looking into the limitations on time. Asking instructors how they felt about expectations and the analytics could identify an impact that was not evident from expectations or patterns in behavior alone. In this thesis, analytic use was treated as an indicator for instructor reflection. Using tools to measure reflection or interviewing instructors before and after use of this learning analytic would more directly measure reflection. For understanding student behaviors, an analysis of meeting expectations vs. performance (grades) could be done. Do students who meet these expectations perform better in the class? This could also help instructors better reflect on whether or not an expectation is necessary or is positively influencing student performance.

Appendix A

Expectation specification form

Please fill out the form as if you were presenting your expectations for a video or group of videos
Not all fields are required, leave blank if it does not apply to you

Watching Expectation

I want my students to watch $\underset{[amount]}{\rule{4cm}{0.4pt}}$ of a video before $\underset{[deadline]}{\rule{4cm}{0.4pt}}$.

Posting Expectations

I want my students to post [1. Top-level posts only, 2. Replies only, 3. (anything/blank)]…

1. I want my students to post $\underset{[how\ many?]}{\rule{2cm}{0.4pt}}$ **top-level posts** of type(s) $\underset{[annotation\ types]}{\rule{4cm}{0.4pt}}$ before $\underset{[deadline]}{\rule{4cm}{0.4pt}}$
2. I want my students to post $\underset{[how\ many?]}{\rule{2cm}{0.4pt}}$ **replies** to $\underset{[Who?\ or\ what\ annotation\ types?]}{\rule{5cm}{0.4pt}}$ before $\underset{[deadline]}{\rule{4cm}{0.4pt}}$.
3. I want my students to post $\underset{[how\ many?]}{\rule{2cm}{0.4pt}}$ $\underset{[annotation\ types]}{\rule{4cm}{0.4pt}}$ before $\underset{[deadline]}{\rule{4cm}{0.4pt}}$.

Quiz expectations

I want my students to answer all of the quiz questions in this video before $\underset{[deadline]}{\rule{4cm}{0.4pt}}$.

On a scale from 1 to 7 (1 is strongly disagree, 7 is strongly agree)
This form was easy to fill out

1 2 3 4 5 6 7

Why did you give it that score?:

This would cover the expectations I have for my classes that use TrACE

1 2 3 4 5 6 7

Why did you give it that score?:

A.1 Instructor Responses

- I want my students to watch *100%* of a video before *posted deadline.*
- I want my students to watch *between 75 and 100%* of a video before *the start of lab.*
- I want my students to post *between 1 and 3* top-level posts of type(s) *anything* before *the start of lab.*
- I want my students to post *1* top-level posts of type(s) *Comment/Question* before *an hour before class.*
- I want my students to post *1* replies to *Comprehension Check* before *an hour before class.*
- I want my students to answer all quiz questions in this video before *an hour before class.*
- I want my students to answer all quiz questions in this video before *midnight before class.*

Instructors reported that the form covered their expectations, reporting that *"It's basically what I already do"* and *"Ensuring vids watched and questions answered by due* *time are my main concerns"*. There were some expectations where instructors wanted to reply to "all" posts of a certain type. For example, replying to all instructor posts or replying to all all Comprehension Check posts in a video. Although this was challenging to write into this form, the prompt was modified and it was successfully implemented in the prototype.

Expectations that could not be coded using this form and were not included in the final prototype involve student groups or expectations that span multiple videos. For example:

- I want students *in the posting group* to create at least 1 post of type question before the start of class.
- I want students *in the reply group* to reply to at least 3 posts made by students in the question-asking group before the start of class.
- I want *discussion leader* students to reply to all questions posted by other students.
- I want all students to post at least 2 times *between these 4 videos* before the deadline.

To consider student groups, a quick way to create groups would need to be created, and it would need to allow for instructors to dynamically create groups. Dividing a class into groups can be an impromptu activity, and doing so within the system should not be cumbersome.

Bibliography

Ali, L., Hatala, M., Gašević, D., and Jovanović, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers & Education*, 58(1):470–489.

Argyris, C. and Schön, D. (1978). *Organizational learning: A theory of action approach.* Addision Wesley, Boston, MA.

Arnold, K. E. (2010). Signals: Applying academic analytics. *EDUCAUSE Quarterly*, 33(1):n1.

Arnold, K. E. and Pistilli, M. D. (2012). Course signals at Purdue: using learning analytics to increase student success. In *Proceedings of the Second international conference on learning analytics and knowledge*, pages 267–270.

Barab, S. and Squire, K. (2004). Design-based research: Putting a stake in the ground. *The journal of the learning sciences*, 13(1):1–14.

Beyer, H. and Holtzblatt, K. (1999). Contextual design. *interactions*, 6(1):32–42.

Bienkowski, M., Feng, M., and Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. Technical report, U.S. Department of Education.

Charleer, S., Santos, J. L., Klerkx, J., and Duval, E. (2014). Improving teacher awareness through activity, badge and content visualizations. In *New Horizons in Web Based Learning*, pages 143–152. Springer International Publishing, Switzerland.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., and Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, 4(5-6):318–331.

Clow, D. (2012). The learning analytics cycle: closing the loop effectively. In *Proceedings of the Second international conference on learning analytics and knowledge*, pages 134–138.

Collins, A., Joseph, D., and Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the learning sciences*, 13(1):15–42.

Dawson, S., Bakharia, A., and Heathcote, E. (2010). SNAPP: Realising the affordances of real-time SNA within networked learning environments. In *Proceedings of the 7th International Conference on Networked Learning*, pages 125–133.

Dennen, V. P., Aubteen Darabi, A., and Smith, L. J. (2007). Instructor–learner interaction in online courses: The relative perceived importance of particular instructor actions on performance and satisfaction. *Distance education*, 28(1):65–79.

Dorn, B., Schroeder, L. B., and Stankiewicz, A. (2015). Piloting TrACE: Exploring spatiotemporal anchored collaboration in asynchronous learning. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 393–403.

Duval, E. (2011). Attention please!: learning analytics for visualization and recommendation. In *Proceedings of the First International Conference on Learning Analytics and Knowledge*, pages 9–17.

Dyckhoff, A. L., Lukarov, V., Muslim, A., Chatti, M. A., and Schroeder, U. (2013). Supporting action research with learning analytics. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 220–229.

Elson, J. (2016). Formative assessment in an online asynchronous learning environment. Master's thesis, University of Nebraska at Omaha.

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5-6):304–317.

Goodyear, P. (2010). Teaching, technology and educational design: The architecture of productive learning environments. *Final report Senior Fellowship, ALTC*.

Greller, W. and Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational technology & society*, 15(3):42–57.

Gulliksen, J., Lantz, A., and Boivie, I. (1999). User centered design in practice-problems and possibilities. Technical Report TRITA-NA-D9813, Department of Computer Science, Royal Institute of Technology.

Jay, J. K. and Johnson, K. L. (2002). Capturing complexity: A typology of reflective practice for teacher education. *Teaching and teacher education*, 18(1):73–85.

Kensing, F. and Blomberg, J. (1998). Participatory design: Issues and concerns. *Computer Supported Cooperative Work (CSCW)*, 7(3-4):167–185.

Lacher, L. L. and Lewis, M. C. (2015). The effectiveness of video quizzes in a flipped class. In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*, SIGCSE '15, pages 224–228.

Larrivee, B. (2008). Development of a tool to assess teachers' level of reflective practice. *Reflective practice*, 9(3):341–360.

Lockyer, L. and Dawson, S. (2012). Where learning analytics meets learning design. In *Proceedings of the Second International Conference on Learning Analytics and Knowledge*, pages 14–15.

Luik, P., Taimalu, M., Kalk, K., Täht, K., et al. (2014). Validity and reliability of two instruments to measure reflection: A confirmatory study. *TRAMES*, 18(2):121–134.

Maher, M. L., Latulipe, C., Lipford, H., and Rorrer, A. (2015). Flipped classroom strategies for CS education. In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*, pages 218–223.

Murphy, D. L. and Ermeling, B. A. (2016). Feedback on reflection: comparing rating-scale and forced-choice formats for measuring and facilitating teacher team reflective practice. *Reflective Practice*, 17(3):317–333.

Muslim, A., Chatti, M. A., Mahapatra, T., and Schroeder, U. (2016). A rule-based indicator definition tool for personalized learning analytics. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*, pages 264–273.

Nelson, T. H., Slavit, D., Perkins, M., and Hathorn, T. (2008). A culture of collaborative inquiry: Learning to develop and support professional learning communities. *Teachers College Record*, 110(6):1269–1303.

Nicol, D. J. and Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education*, 31(2):199–218.

Romero, C., Romero, J. R., Luna, J. M., and Ventura, S. (2010). Mining rare association rules from e-learning data. In *Educational Data Mining 2010*, pages 171–180.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science*, 18(2):119–144.

Scheffel, M., Drachsler, H., Stoyanov, S., and Specht, M. (2014). Quality indicators for learning analytics. *Journal of Educational Technology & Society*, 17(4):117–132.

Scheffel, M., Niemann, K., Pardo, A., Leony, D., Friedrich, M., Schmidt, K., Wolpers, M., and Kloos, C. D. (2011). Usage pattern recognition in student activities. In *Towards Ubiquitous Learning*, pages 341–355. Springer.

Schön, D. (1987). *Educating the reflective practitioner*. Josey-Bass Publishers, San Francisco, CA.

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

Sparks-Langer, G. M., Simmons, J. M., Pasch, M., Colton, A., and Starko, A. (1990). Reflective pedagogical thinking: How can we promote it and measure it? *Journal of teacher education*, 41(5):23–32.

Van Manen, M. (1995). On the epistemology of reflective practice. *Teachers and Teaching: theory and practice*, 1(1):33–50.

Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, pages 203–211.