RESEARCH REPORT: INSTITUTION A

Findings from an administration of the ISOP framework at Institution A in the Spring of 2013:

Insights into course planning, classroom teaching, and student experiences in STEM courses

Matthew T. Hora, Jana Bouwma-Gearhart, Amanda Oleson, and Jennifer Collins[1]

Tracking the processes of data-driven decision making in higher education (TPDM)

(NSF # DUE-1224624)

JULY 2013

S U M M A R Y

In the spring of 2013, our research team spent a week at Institution A collecting data on teaching and learning in science, technology, engineering, and mathematics (STEM) disciplines. In light of increasing efforts to improve the quality of undergraduate education by using data to inform and assess curriculum and instruction, our goal is to document and describe faculty practices in these areas in order to identify if and how data are being used to enhance their current practices. To capture the various features of the teaching and learning dynamic, we used the Instructional Systems of Practice (ISOP) framework that includes measurements of (1) faculty course planning, (2) classroom teaching and (3) student perceptions of teaching quality and study habits. This report provides a summary of results from interviews (n=20 faculty), classroom observations (n=20 courses), student surveys (n=13 courses), and student focus groups (n=6 groups) from undergraduate courses in mechanical engineering, geoscience, physics, and biology. We also report on the alignment of these datasets for a sub-sample of 12 courses. Conclusions focus on the types of planning and teaching activities observed at Institution A, and recommendations are provided in regard to educational improvement initiatives.

---

[1] Hora and Oleson are affiliated with UW-Madison, and Bouwma-Gearhart and Collins are affiliated with Oregon State University.

[2] In this paper the terms "faculty" and "instructor" are used interchangeably to denote individuals who are the lead teachers for an undergraduate course. These individuals include both tenure-track research faculty, full-time lecturers, tenure track teaching faculty, and others designated as lead instructors.

[3] These figures vary because the number of online surveys respondents were able to send to their courses was less than the total number of courses included in the study. Similarly, the number of focus groups that the research team could manage in a week's time was less than the total number of courses.

Some of the notable results from the data include the following:

Faculty course planning

- Planning procedures for courses are generally formal and structured when creating a new course, but more ad-hoc when revising existing courses
- Less than 50% of the faculty respondents reported that their courses were evaluated at the end of the term, besides end-of-term student evaluations
- Faculty report often using data for formative and summative purposes in their courses
- Faculty report expecting students to be actively thinking during class (n=12), following along (n=6) and actively questioning instructors (n=5)

Classroom teaching

- Students were observed actively speaking or actively engaging in the class in 55% across all two-minute intervals that comprised a class period
- Faculty were observed lecturing with PowerPoints in 63% of all two-minute intervals, lecturing at the chalkboard in 26%, and lecturing with demonstrations in 5%
- Students were asked questions seeking original information in 42% of all two-minute intervals
- The most common form of cognitive engagement provided to students was receiving and memorizing information (37%), followed by making connections to the real world (24%) and problem-solving (20%)

Student perceptions and study habits

- Students prefer a "surface" or transmitting information approach to teaching over a "deep" or supporting understanding approach (n=248)
- Students report having a "surface" and a "deep" approach to learning in equal measure (n=256)
- Students report spending between 6-10 hours a week studying for all of their courses, and between 1-2 hours for the "target course" being studied here (n=256)

Background to the Study

In the spring of 2013, a team of researchers from the University of Wisconsin-Madison (UW-Madison) and Oregon State University (OSU) conducted interviews with and classroom observations of STEM faculty, as well as conducted focus groups and administered surveys to students at Institution A. These efforts were part of a study focusing on the nature of curriculum design in STEM departments and subsequent influences on teaching and learning. [footnote: This study is supported by the National Science Foundation (Award DUE-1224624).]

This study is informed by research on data-driven decision making (DDDM) in K-12 schools, which is a widely advocated approach by policymakers and researchers that is based on the use of rigorous data (e.g., regarding student achievement, teacher evaluations) to inform school and district-wide decisions, as opposed to ad-hoc and anecdote based decision-making procedures. The recent push for the use of educational data in the U.S. was largely sparked by the passage of the No Child Left Behind (NCLB) legislation in 2002. However, issues with DDDM include challenges with translating mountains of data into actionable knowledge, the lack of local expertise or systems for collecting and interpreting data, and the focus on using data to measure teacher "quality." In essence, data has become a four-letter word in many educational circles, although some observers note that a "paradigm shift" is taking place that emphasizes providing data to help educators rather than towards punitively assessing them (Mandinach, 2012). This shift entails moving away from a sole focus on accountability to an approach that is more focused on supporting teachers' growth and development in ways that build upon and respect their professional knowledge. This stance also encourages designing new policies and initiatives that build upon a group's existing practices and expertise, rather than administering change in a top-down manner that ignores local practice and cultural realities (Spillane, Halverson & Diamond, 2001).

While no equivalent of NCLB exists at the postsecondary level in the U.S. or Canada at the present time, interest in educational improvement at the undergraduate level is increasing. A long-standing critique of the quality of undergraduate education (e.g., Arum & Roksa, 2012; Bok, 2002) has more recently been bolstered by a specific focus on teaching and learning in the STEM disciplines, based on their perceived centrality to growth and maintaining a competitive edge in a global economy, as well as a scientifically literate population. In addition, accreditation agencies are increasingly requiring institutions to demonstrate that they are collecting teaching and learning-related data. With the growing emphasis on educational improvement and accountability in postsecondary settings, we wished to better understand the "state of the art" of data use for curriculum and instruction in STEM departments. Importantly, the motivation for documenting these behaviors is not to critique current practice, but instead to provide educational leaders, policymakers, and faculty with an empirically grounded account of teaching practices and data use that can be used as a starting point for thinking about enhancing existing DDDM or designing new DDDM models. Insights into these practices will shed light on the degree to which departments and faculty are actively engaged in rigorous continuous improvement efforts; which can be seen as a precursor to the adoption or consideration of inquiry-based teaching or sophisticated course design.

The motivation for documenting these behaviors is not to critique current practice, but instead to provide educational leaders, policymakers, and faculty with an empirically grounded account of teaching practices and data use that can be used as a starting point for thinking about enhancing existing DDDM or designing new DDDM models.

We also are conducting a field experiment by providing new data about teaching and learning (this report) to curriculum designers at program/department and course levels to see if these data are at all useful. The hypothesis is that reflective practice is essential for good teaching, and that the ubiquitous student evaluations and infrequent peer observations are insufficiently detailed to spark instructors' reflection on their own teaching, or to catalyze administrators' reflections on their own program or department. Thus, we offer this alternative set of data about teaching and learning, and we are actively seeking feedback about their utility and how the dataset itself and/or the reporting of it can be improved in the future.

Theoretical Framework: Instructional Systems-of-Practice

A core theoretical proposition guiding our study is that teaching is best understood as a multi-dimensional phenomenon in which various elements interact with one another over time, both within the classroom and across multiple learning environments. We contend that the common approach of conceptualizing postsecondary teaching solely as the regularity with which certain teaching methods are used in the classroom (e.g., lecturing, small group work) is an overly reductionist and de-contextualized approach that obscures the complexity of both classroom practice and teaching in general. This stance is partially an artifact of the reliance on questionnaires that are commonly used to study postsecondary teaching. We argue that a more appropriate way to measure the complex phenomenon and in-depth nuances of classroom teaching is via in-person observations.

However, even a detailed observation only tells part of the story. To ascertain why educators do what they do, one must actually talk to them. Further, to understand how the classroom experience is actually influencing students, it is necessary to also collect data from students. As part of a previous study (the CCHER project based at UW-Madison) the Instructional Systems of Practice (ISOP) framework was created in order to capture the multi-faceted components of teaching and learning in a comprehensive fashion. The framework is based on systems-of-practice theory from distributed cognition that emphasizes how behavior is best understood in terms of the interactions among actors, artifacts, and features of the task itself (Halverson, 2003; Hora & Ferrare, 2012). The ISOP framework is comprised of three primary elements: course planning, classroom instruction, and student classroom experiences. Importantly, the framework does not evaluate student learning or evaluate teaching quality, but instead is intended as a way to describe critical features of teaching and learning in a rigorous fashion.

The framework does not evaluate student learning or evaluate teaching quality, but instead is intended as a way to describe critical features of teaching and learning in a rigorous fashion.

Methods

Research universities were selected for this study in part because of the large number of undergraduates being trained in STEM disciplines at these institutions. The three study sites had similar undergraduate enrollments, external research funding, and various STEM pedagogical improvement initiatives underway at the time of data collection. The disciplines included in this study are physics, biology, geoscience, and mechanical engineering. The course component of interest in this study was the class lesson, popularly known as the "lecture" or "recitation." That is, laboratory, discussion sections, tutorials, field-work, and online components of courses were not observed. The instructors self-selected into the study, and thus the results should not be generalized to the larger population of faculty at these institutions. Most courses observed in this study had between 26 and 200 students, were at or below the 300 level, and were taught by either full professors or non tenure-track instructors. For more information about the study sample see Appendix A. Importantly, each of the datasets pertain to a specific "target course," such that a single course (e.g., Physics 200) was the topic of the faculty interviews, classroom observations, student focus groups, and student surveys.

Four types of data were collected that reflect the different components of the ISOP framework. First, a semi-structured interview with faculty focused on eliciting their goals for a specific class and their expectations for their students while in class. These interviews were recorded and analyzed for recurring themes and patterns regarding the research questions. Second, classroom observations were conducted with faculty using the Teaching Dimensions Observation Protocol (TDOP) instrument, which captures five dimensions of instruction in two-minute intervals. Third, an online survey was administered to students that focused on their perceptions of teaching quality for the observed class, their preferences for teaching, and their study habits. Fourth, a focus group was held with a sub-sample of courses where students were asked to discuss their views of teaching and their approaches to studying. Prior to gathering data in the field, the four researchers underwent a rigorous training. In particular, TDOP required considerable training, and an acceptable inter-rater reliability (IRR) was established at the conclusion of the training process. For more information about IRR see Appendix B, and all research instruments are available at the project website (tpdm.wceruw.org).

RESULTS: Curriculum design and use of data in planning

In this section we report the results from the first portion of the ISOP framework: how faculty and administrators plan undergraduate courses. In particular, we focus on the nature of the planning processes (i.e., ad-hoc or structured), the role of data in planning, and the existence of course assessments and "feedback loops" for continual improvement of courses. Results are presented for these key topics in Table 1 below.

TABLE 1: ***Interview responses regarding the use of data for curriculum design***

	All n= 19	Biology n= 9	Mech Eng n= 3	Geology n= 4	Physics n= 3
Planning procedures					
Ad-hoc and unstructured	15	7	1	3	3
Formal and structured	13	7	2	2	2
Learning objectives					
Learning objectives	10	6	1	3	0
Learning objectives linked to data	8	6	0	2	0
Use of data for planning					
Formative purposes	14	9	3	2	0
Summative purposes	14	8	3	2	1
Is the course itself evaluated?	8	5	1	2	0
Feedback loop exist for continuous improvement?	14	8	3	3	0

Planning procedures: Degree of formality and structure

One of the primary research questions in our study is whether or not curriculum planning procedures in academic departments are formal or informal. The results indicate that 15 respondents reported that planning was ad-hoc and unstructured, and 13 reported that planning was a structured affair. These data appear contradictory until it is made clear that some respondents answered differently depending upon the level of analysis being considered. That is, the degree of formality and structure largely depends on the administrative level being considered. At the department or program level, there tends to be more structure in terms of creating or altering courses, and adding new courses to degree requirements. Planning is also formalized in cases where faculty desire to create a new course, which requires approval at the College level, along with the satisfaction of numerous requirements and committees.

In contrast, at the course level there tends to be much more informality to the curriculum planning process. In these cases, respondents reported that they enjoyed a high degree of autonomy when changing their own courses, and that no policies or departmental structures existed that governed what and how they teach. One exception includes introductory courses that are taught across multiple sections, which necessitates a certain degree of formality and structure. In future analyses of the data, we will tease apart these distinctions more carefully. That being said, some respondents did feel that the planning process regarding course content was structured, whereas the planning process regarding pedagogy was mostly ad-hoc at both the program and course levels. This suggests that a possible distinction can be made between the curriculum as a list of topics (i.e., as reflected by the syllabus) and the actual instruction of these topics in the classroom or laboratory.

Learning objectives

Research on curriculum planning emphasizes the role of learning objectives as a way for instructors to align their content covered, classroom teaching, and assessment materials. Given the focus on data use in this study, we are also interested in how learning objectives influence data use for teaching and learning purposes. Among the study sample, 10 respondents reported that they had learning objectives for their courses, with eight respondents linking these objectives to data collected in the course. These data were typically in the form of exams and quizzes. Several faculty observed the influence of a pedagogical reform initiative in the articulation of learning objectives, though some faculty had instituted them with little or no interaction with the initiative.

Use of formative and summative data for planning

In interviews with faculty, we asked about the use of data for designing and/or revising a course. Overall, 14 respondents reported the use of data for formative planning and revision of the curriculum. This particular use of data involved reviewing things such as online quizzes, homework assignments, and responses to clicker questions in ways that led to the updating or revision of the course. Importantly, this use of data involved a period of reflection where the respondent takes the time to think about the implications of the data for the course. In regard to the use of data for summative purposes, 14 respondents reported that they used data collected at the end of the course in order to update or revise the course for the next term. In these cases, final projects or exams, end-of-term student evaluations of the course, and conversations with students constituted the information that faculty used in their planning. Student evaluations were by far the most commonly collected and utilized type of data for summative purposes.

It is important to note that our conception of the term data is rather broad. Interestingly, several faculty requested clarification regarding the term "data," as it was a term that was not often associated with teaching. Instead, data was seen as something used in their disciplinary research. Further, while our own definition of data is rather broad and extends beyond the common perception of data as numeric figures stored or tabulated in a structured manner, respondents most often viewed data as these things. As such, it was not uncommon for faculty to respond "no" to this question, and then list a number of sources from which they drew information about the curriculum that we DO in fact consider data. Examples of this include textbooks, prior instructors' syllabi and PowerPoint slides, and even informal conversations with students during office hours. This broad view of what constitutes data is used in the organizational learning literature and will be explored in greater detail as part of this study.

Course evaluation and assessment

We also asked faculty if their course was evaluated or assessed in some fashion. A common response was "Do you mean am I evaluated or is the course evaluated?" This distinction was understandable given the ubiquity of student evaluations, which for several respondents doubled as an evaluation of their own teaching quality and of the course itself. In any case, eight respondents answered, "yes" to this question, and in most cases described an informal process for assessing whether or not the course was deemed "successful." These included end-of-term meetings among faculty directly involved in teaching a course or more commonly, student evaluations. More sophisticated examples included pre- and post-testing of student learning, and these cases were all associated with a pedagogical reform initiative. In no instances were any formal procedures in place at the departmental level to ensure that some sort of course evaluation process would take place beyond the administration of end-of-term student evaluation surveys. Thus, the assessment of whether or not a course was successful was left up to the individual instructor or team of instructors, and to students via their end-of-term evaluations.

Feedback for continuous improvement

Related to the question about the use of data for planning, faculty were asked about whether or not there existed any mechanisms for "feeding" data into systems for evaluating or assessing courses. This question was motivated by a desire to ascertain if institutions, departments, and individual faculty were actively engaged in efforts to continually improve their courses. In management and organization science this is known as "continuous improvement," and researchers encourage the use of robust evidence as part of these systems that are essentially DDDM systems. Fourteen faculty reported that some sort of feedback loop did exist for the continual improvement of their courses, with several being based on regular and/or end-of-term meetings with other faculty. These faculty were either involved in teaching other sections of a multiple section course, were team-teaching a course, or simply had interested colleagues. However, in only a few cases did these meetings involve a rigorous analysis of evidence to assess course efficacy or quality; instead, they were described as relatively informal affairs where successes and challenges from the previous term were discussed. That said, the fact that the meetings were held at all is commendable and could be considered a foundation for building future continuous improvement systems.

Other findings

Finally, faculty described some of the factors that influence the curricular decision-making process. A preliminary review of the data indicates that some of these factors include accreditation policies, a pedagogical reform initiative, textbooks, consideration of learning goals, and a focus on conveying the canon of their discipline.

RESULTS: Focus on faculty expectations for student cognition during class

One of the foci of the ISOP framework in regard to faculty planning is the instructional goals and expectations for students that teachers have for their students. In particular, we emphasize the expectations for student cognitive activity during the class period. While the focus on student thinking or cognition varies from questions pertaining to instructional goals, the resulting answers are similar to previous studies conducted by our research group that elicited information about goals. The focus on instructors' expectations for cognitive activity is based on the desire to explore (a) the ways in which faculty conceptualize student thinking while in the classroom, and (b) to explore the alignment between faculty expectations for cognitive activity and the classroom observation data. It is important to note that some respondents cited more than one type of expectation (see Table 2, below).

TABLE 2: ***Interview responses regarding expectations for students' mental activity in class***

	All n= 19	Biology n= 9	Mech Eng n= 3	Geology n= 4	Physics n= 3
Expectations for student mental activity in class					
Active engagement (active thinking/critical thinking)	12	7	1	4	0
Follow along/be attentive	6	3	3	0	0
Active questioning	5	2	2	1	0
Come to class prepared to think/engage	4	4	0	0	0
Active engagement (hands-on/direct application)	3	3	0	0	0
Struggling with material	3	1	0	1	1
Talk and interact with peers	2	0	0	1	1
Understand the material	2	0	1	0	1
Integrate information across topics	2	1	0	0	1
Stay awake	1	1	0	0	0

One of the notable features of these data is that the categories for student cognition reported by faculty do not align with the categories regularly used in the literature. These include categories from Bloom's taxonomy, commonly used surveys such as the National Survey of Student Engagement (NSSE), and even the cognitive engagement codes from our own TDOP classroom observation instrument. Examples of these categories include "evaluate material and consider alternative explanations" or "synthesize information from disparate sources." While the responses of faculty are not completely dissimilar from constructs such as these, the results suggest that some faculty in the study sample instead think of student engagement in less specific terms. The four most commonly reported expectations are briefly described in the following text.

Active engagement: active thinking/critical thinking

The most commonly reported expectation (12 respondents) was that students be actively thinking during class such that they were engaged with the material in a deep and critical manner. One aspect of this type of engagement was described as "critical thinking," where students were not simply taking notes or memorizing facts, but actively critiquing the ideas being presented. In this way, the expectation was similar to a less reported one of students actively "struggling" with the material.

Follow along/be attentive

The next most common expectation (six respondents) was that students be attentive and follow the lecture during the class period. This expectation was described in terms that suggested lower expectations for students than active thinking. Here, students are simply expected to be paying attention to the instructor.

Active questioning

Five faculty reported the expectation that students be actively questioning the material and themselves during class. This expectation encompasses the view that students should interrupt instructors with questions when they do not understand the material. This expectation is closely related to active thinking in that represents an expectation that students are not simply paying attention, but will be active participants in the classroom experience, whether through mental activity or verbal questioning.

Come to class prepared to think and be engaged

Finally, four respondents expressed an expectation that students come to class adequately prepared and ready to be actively engaged in the class.

These results indicate that for the faculty in the study sample, expectations for student thinking can be categorized as a relatively high expectation for active thinking and questioning (and the attendant expectation that students come to class prepared to be actively engaged), and a lower expectation that students are merely following along or paying attention.

RESULTS: Classroom teaching

In this section we present data from classroom observations of 19 STEM faculty. A summary of selected TDOP codes are provided in Table 3, below. In order to aid in the interpretation of these data, we also analyze certain codes and code-combinations. It is important to note that we do not claim that the TDOP is on its own measuring either teaching quality or student learning outcomes – to do so in a robust manner requires insights into course design, out-of-class learning opportunities, and of course, evidence about student learning outcomes. However, we suggest that certain types of pedagogical behaviors may be associated with effective teaching, and that these behaviors can "set the stage" for student learning.

We do not claim that the TDOP is on its own measuring either teaching quality or student learning outcomes – to do so in a robust manner requires insights into course design, out-of-class learning opportunities, and of course, evidence about student learning outcomes

TABLE 3: ***Classroom observation data using the Teaching Dimensions Observation Protocol (TDOP)***

	All n= 19	Biology n= 9	Mech Eng n= 3	Geology n= 4	Physics n= 3
Teaching Methods					
Lecturing	.03	.03	.08	.00	.02
Lecturing w/pre-made visuals (e.g., PowerPoint)	.63	.87	.33	.40	.56
Lecturing w/hand-made visuals (e.g., chalkboard)	.26	.10	.66	.29	.23
Lecturing w/demonstration	.05	.02	.15	.00	.07
Interactive lecture (i.e., 2+ questions posed)	.03	.03	.06	.00	.02
Working through problems	.06	.03	.15	.00	.10
Small group work	.16	.12	.00	.38	.15
Desk work	.03	.02	.00	.01	.09
Multi-media	.02	.02	.03	.00	.03
Pedagogical Moves					
Humor	.14	.22	.12	.04	.02
Illustration or anecdote	.20	.18	.21	.25	.11
Graphic	.53	.65	.39	.46	.46
Organization	.12	.13	.09	.12	.11
Emphasis	.07	.08	.04	.04	.07
Assessment	.12	.17	.01	.09	.11
Teacher-Student Interactions					
Rhetorical questions	.12	.13	.03	.15	.14
Display questions (i.e., seeking new information)	.42	.48	.38	.40	.31
Comprehension questions	.09	.09	.12	.08	.09
Student novel question	.02	.02	.03	.00	.02
Student comprehension question	.12	.10	.23	.13	.07
Student response to teacher question	.34	.38	.31	.35	.21
Student peer interactions	.15	.09	.01	.32	.24

	All n= 19	Biology n= 9	Mech Eng n= 3	Geology n= 4	Physics n= 3
Cognitive Engagement					
Receiving and memorizing	.37	.46	.33	.26	.28
Problem-solving	.20	.15	.01	.46	.27
Creating	.04	.01	.00	.15	.00
Connecting to real-world	.24	.20	.23	.38	.11
Student Engagement					
Very high	.70	.74	.64	.89	.32
High	.17	.21	.21	.04	.19
Medium	.09	.04	.11	.06	.29
Low	.01	.00	.03	.00	.04
Instructional Technology					
Chalkboard	.12	.00	.33	.29	.00
Overhead projector	.10	.03	.39	.00	.10
PowerPoint	.58	.88	.05	.39	.55
Clickers	.12	.17	.01	.09	.12
Demonstrations	.03	.00	.11	.00	.04
Digital tablet	.11	.11	.10	.00	.27
Movies or simulations	.02	.03	.01	.01	.03
Student Talk					
DW/SGW/SR/PI/SNQ/SCQ	.55	.52	.53	.73	.43

Note: these data reflect the proportion of all 2-minute intervals where a particular code was observed.

Prevalence of lecturing

One of the notable findings is the prevalence of lecturing in the observed classrooms. In the TDOP, lecturing is conceptualized in five different ways: stand-alone lecture where the instructor is speaking to the class with no instructional technology or interactions with students (L), lecturing with a pre-made visual such as PowerPoints (LPV), lecturing with hand-written visuals such as board work (LHV), lecturing while also performing a demonstration (LDEM), and interactive lecturing (LINT) where the instructor is actively engaged with students in a Socratic dialogue. Overall we observed high rates of lecturing with pre-made visuals (63% of all two-minute intervals), and lecturing with hand-made visuals (26%), with some interesting variation across disciplines. It is important to note however that lecturing, which is essentially an instructor speaking to a class, is on its own not an indicator of instructional quality. When used in a pedagogically deliberate manner and/or in conjunction with other teaching methods or pedagogical moves, lecturing can be an important and valuable component of teaching (Saroyan & Snell, 1997).

Pedagogical clarity and organization

Research on effective teaching in postsecondary classrooms demonstrates that teacher clarity and the quality of a lesson's organization are strongly related to student outcomes (Murray, 1999). These studies largely draw on multi-item survey scales that students fill out at the conclusion of a class or semester. The TDOP necessarily simplifies these more nuanced measures into single codes, which provide a rough snapshot of an individual's pedagogical clarity and organizational skills within a class. Among the study sample, faculty were observed in

7% of all two-minute intervals using behaviors associated with clarity (e.g., emphasizing important points for students), and in 12% of all intervals using explicit verbal markers for the lesson's organization (e.g., marking transition points between topics).

Making connections to students' lives

One of the hallmarks of effective lecturing and public presentations is the use of anecdotes and details that make connections between the topic being discussed and the daily lives and experiences of an audience. The TDOP captures the pedagogical strategy of illustration (IL) use, which is always co-coded with the cognitive engagement of making connections to the real world (CN). In the study sample, IL was observed in 20% of the two-minute intervals, and CN was observed in 24%.

Peer-to-peer learning

We observed small group work being used in 16% of all two-minute intervals, with the largest proportion being observed in the geology group (38%). Another indicator of peer-to-peer learning is captured in the peer interactions (PI) code under student-teacher interactions. We observed PI in 15% of the two-minute intervals, with high rates in the geology group (32%) and the physics group (24%). One would expect to see small group work and peer interactions at similar rates of incidence (e.g., geology group), but in some cases students may be interacting without clear instructions to break up into pairs or small groups – this may be the explanation for the discrepancy between the two for the physics group (who had 15% of intervals coded as small group work).

Socratic lecturing

The Socratic method is a pedagogical approach that involves a dialogue between teacher and students, where the teacher poses a series of probing questions in a way that explores the beliefs, assumptions, and knowledge of students. In the TDOP, we have operationalized this type of teaching as interactive lecture (LINT), where the instructor asks a series of two or more related questions to the class, while allowing student answers to help dictate the course of the discussion. This can involve a question such as "What is the difference between a bacteria and a virus?" with the instructor taking answers and elaborating upon each in turn, and allowing responses to build upon one another. For the study sample, we observed interactive lecturing in 3% of all two-minute intervals.

In-class assessment

In-class assessment is one way that instructors can determine how well (or not well) students understand the material being taught. Insights into student misconceptions or lack of comprehension can provide valuable information for an instructor regarding whether or not students have understood the material and if further instruction is necessary. These assessments can either be formative or summative in nature and can take many forms, including traditional quizzes and tests, online quizzes, in-class questions, and so on. Three of the TDOP codes capture forms of formative assessment: the generalized assessment code which includes in-class quizzes, and other explicit assessment codes; clicker questions; and display questions (e.g., What is ecological equilibrium?). Across the study sample, some form of assessment was observed in 12% of the two-minute intervals; clickers were observed in 12% of the intervals; and display questions were observed in 42% of the intervals.

Student engagement

One of the biggest concerns of instructors is whether or not their students are paying attention in class, and ideally, are actively engaged with the material and pedagogy. This is a very difficult phenomenon to measure unless students are self-reporting their state of mind in real-time. With the TDOP, we use a very coarse measure to determine student engagement, and include these data more for informational purposes rather than as a robust measure of student engagement. Part of the reason for these caveats is the difficulty in collecting data on student engagement. Data are collected during the observations by taking a sort of "transect" and identifying 8-12 students sitting nearby the observer (positioned in various locations in classes). These students are observed every two minutes and are considered "engaged" if they are not talking, surfing the Internet, sleeping, reading, or looking away from the instructor. Thus, our measures of students being engaged are simply if they are looking at the instructor or otherwise doing a class activity (such as taking notes or working on a problem as prompted by the instructor) with no visible distraction. Nothing is known about what is actually going through their minds and/or the quality of their attention. Across the study sample, in 70% of all two-minute intervals students in the immediate vicinity were considered "very highly engaged" (i.e., more than 75% of the students were paying attention), 17% were considered "highly engaged) (i.e., more than 50%), 19% were considered "medium" engaged (i.e., between 25-50%), and in 1% "low" engagement was observed (i.e., between 0-25%).

Time provided for student "talk"

The amount of time that students are given to vocalize their thoughts and questions is considered to be an important indicator of a classroom where students are actively engaged with the material. For example, Amidon and Flanders (1979) focused on the interactions between students and teachers as representing a key aspect of effective classrooms. Their classroom observation protocol distinguishes between two types of "talk" in the classroom: (a) teacher talk which is either direct (e.g., giving directions) or indirect (e.g., praising, asking questions), and (b) student talk which is considered either as a "response" (i.e., convergent answers to posed questions) or an "initiation" (i.e., divergent questions, or responses to posed questions that depart from the flow of the conversation). While our protocol does not exactly align with this framework, the TDOP does provide insights into the amount of class time that is devoted to student and teacher "talk."

One way to measure the amount of time student(s) are actively engaged during the class period is to assess how much time they are spending either speaking in class and/or involved in activities where they are actively interacting with one another or the material. To capture the amount of class time allocated to "student talk," we identified the percentage of two-minute intervals that included one of the following TDOP codes: desk work (DW), small group work (SGW), student response to questions (SR), peer interactions (PI), student novel questions (SNQ), and student comprehension questions (SCQ). These codes represent a broad range of student behaviors and thus are a coarse measure of "student talk," but we felt it an informative way to interpret TDOP data to provide insights into the relative amount of time during a class where students are not passively listening to the instructor.

The data indicate that for the entire sample students were observed responding to teacher questions in 34% of all two-minute intervals. Regarding student initiated questioning, novel or original questions were observed in 2% of all intervals, and comprehension questions (e.g., What does X mean?) were observed in 12% of all intervals. Another form of student talk is peer-to-peer interactions, where students are directly engaged with one another without the instructor guiding the class period. This form of student talk was observed in 15% of all intervals.

RESULTS: Student perceptions of teaching

In this section we report descriptive statistics from the survey administered to students immediately after the classes observed with the TDOP. Students took the survey online after being sent an email from the instructor of the course. Students in 13 courses took the survey and the overall response rate was 13.8%, which is extremely low. Research suggests that response rates to online surveys in postsecondary settings are becoming problematic, as students, faculty and staff are inundated with electronic communications (Umbach, 2004). **As a result of this low response rate, the descriptive results reported here should be interpreted with caution and viewed as incomplete accounts of student experiences within the sampled courses.** This is because these data are an incomplete reflection of what students are thinking and doing in the institution's classrooms. With this caveat in mind, these data do still capture the experiences of approximately 275 students and thus may provide an interesting and informative snapshot into the undergraduate experience, particularly as it pertains to their approaches to learning and studying.

In Table 4, descriptive data (means and standard deviations) are provided about student perceptions of teaching at the institution and disciplinary level. Note that the sample size varies according to each of the survey items. Also, each of the responses pertains to a specific class period that was also observed using the TDOP. Information about scale reliability is provided in the Appendix C, and references for the original survey instruments are provided in the resources section at the end of this report. Descriptions of these items and scales are discussed in detail in the following text.

TABLE 4: ***Results from survey of students' perceptions of teaching***

	Institution-Wide (13 courses)		Biological Sciences (6 courses)		Physics (2 courses)		Earth/Space Sciences (4 courses)		Mechanical Engineering (1 course)	
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
Perceptions of teaching quality										
(1=never, 5=very often)										
Clarity	282	3.81 (.77)	154	3.75 (.71)	76	4.03 (.83)	35	3.62 (.74)	17	3.79 (.96)
Organization	282	4.09 (.74)	154	4.12 (.62)	76	4.16 (.81)	35	3.95 (.86)	17	3.87 (1.09)
Challenge/faculty expectations	282	3.01 (.75)	154	3.07 (.73)	76	2.96 (.74)	35	2.98 (.86)	17	2.76 (.77)
Preferences for teaching										
(1=not at all, 5=extremely)										
Supporting understanding (deep)	248	3.07 (.74)	137	3.11 (.73)	68	3.06 (.75)	28	2.85 (.87)	15	3.15 (.57)
Transmitting information (surface)	248	3.69 (.91)	137	3.64 (.90)	68	3.86 (.82)	28	3.49 (1.19)	15	3.78 (.67)
Cognitive activity in classroom										
(1=very little, 4=very much)										
Remember facts (TDOP: RRI)	268	2.68 (.84)	147	2.76 (.80)	71	2.68 (.75)	33	2.48 (1.03)	17	2.47 (1.01)
Describe, explain	268	2.03 (.93)	147	2.05 (.92)	71	1.89 (.92)	33	2.24 (1.03)	17	2.00 (.94)
Apply theories (TDOP: PS)	268	2.62 (.91)	147	2.67 (.85)	71	2.69 (.90)	33	2.18 (1.04)	17	2.65 (.93)
Break into parts	268	2.42 (.90)	147	2.40 (.92)	71	2.49 (.79)	33	2.24 (1.12)	17	2.65 (.61)
Create new ideas (TDOP: CR)	268	2.16 (.93)	147	2.14 (.91)	71	2.24 (.96)	33	2.03 (.88)	17	2.35 (1.00)
Evaluate	268	2.19 (.97)	147	2.22 (.97)	71	2.08 (.95)	33	2.21 (1.02)	17	2.24 (.90)
Make connections (TDOP:CN)	268	2.42 (.94)	147	2.48 (.92)	71	2.24 (1.01)	33	2.36 (.86)	17	2.76 (.97)
Autonomous learning env.										
(1=not at all, 5=extremely)										
Autonomous learning env scale	248	3.43 (.95)	137	3.32 (.90)	68	3.57 (1.00)	28	3.35 (1.03)	15	3.91 (.91)

Perceptions of teaching quality

How students perceive the quality of their instructors' teaching is a widely used method for measuring classroom practice (see Bill & Melinda Gates Foundation, 2012; Wabash National Study of Liberal Arts Education, 2012). Studies concerning student perceptions of teaching differ from typical end-of-term student evaluations in that they often measure a specific feature of a given day's instruction, such as an instructor's skills in organizing and delivering a class. Thus, these measures move beyond global indicators that may strain students' ability to recall instructional practice and are also not vague indicators of student satisfaction or popularity.[4] In this study we used scales from the Wabash National Study of Liberal Arts Education study to capture three important features of instruction that the literature suggests are associated with both effective instruction and student learning (Murray, 1999; Perry & Smart, 1997): clarity, organization, and academic challenge. Results for the entire study sample who answered these questions (n=282) indicate the following (on a scale of 1 being never and 5 being very often):

- Students reported that their instructors' ability to craft organized and coherent lessons was observed with a high degree of regularity (M= 4.09, SD=.74).
- The next highest ranked teaching practice was that of clarity or the ability to present a coherent and understandable class (M=3.81, SD=.77).
- Finally, students reported that their instructors had high expectations and provided substantial academic challenges to them some of the time (M=3.01, SD=.75).

Preferences for teaching

In addition to direct observations of classroom teaching, it is also important to determine how students in a given class are responding to a particular type of instruction because student preferences for teaching are indicative of certain approaches to studying and learning. While we do not suggest that the curriculum and teaching practices should be determined by these preferences, it is undeniable that how students view their learning and the role that the classroom experience plays in this learning is an important part of the teaching and learning dynamic.

For this study we used scales from the Approaches and Study Skills Inventory for Students (ASSIST). The ASSIST survey was developed by Tait, Entwistle and McCune (1998) as part of a long-standing research program on student approaches to studying and learning (e.g., Entwistle & Tait, 1990). A fundamental idea underlying the ASSIST survey is that student approaches to studying and learning exists on a continuum from "deep" (e.g., uses evidence, seeks meaning, interested in ideas) to "surface" (e.g., lacks purpose, bound by syllabus, memorization). These ideas inform the scale for preferences for teaching, which include a sub-scale for supporting understanding (related to a deep approach to studying) and transmitting information (related to a surface approach to studying). Example items include "I prefer lecturers who encourage us to think for ourselves and show us how they themselves think" (supporting understanding/deep approach) and "I prefer

[4] It is important to note that there is a debate in the field about the ability of these surveys to serve as proxy measures for student outcomes, which is a claim made by the Gates Foundation Measuring Effective Teaching Project (see Guarino & Stacy, 2012 and Camburn, 2012 for critiques). In this study, we do not make this claim and instead suggest that collecting student perspectives on teaching efficacy is an important complement to faculty self-reports and classroom observations.

lecturers who tell us exactly what to put down in our notes" (transmitting information/surface approach). Results for the entire study sample who answered these questions (n=248) indicate the following (on a scale of 1 being not at all and 5 being extremely):

- Students reported "somewhat" agreeing with statements that indicated a preference for teaching that supports understanding (i.e., a deep approach) (M=3.07, SD=.74).
- Students reported between "very" and "somewhat" agreeing with statements that indicated a preference for teaching that supports transmitting information (i.e., a surface approach) (M=3.69, SD=.91).

Cognitive activity in the classroom

Next, we asked students to report the types of thinking or cognitive activity that they experience while sitting in the classroom. To identify appropriate items to capture this construct we reviewed existing surveys such as the NSSE, and we conducted cognitive interviews with students at UW-Madison to determine if existing items or conceptualizations of student cognition (e.g., Bloom's taxonomy) made sense to them. In most cases they did not, so we adapted the wording from existing survey items to better fit students natural perception of their classroom cognition and to align with some of the TDOP cognitive engagement measures. The items reported in Table 4 include the following types of cognitive activity (on a scale of 1 being very little, and 4 being very much):

- Remembering facts: Students across the study sample who answered these items (n=268) reported this cognitive activity occurred with relative frequency (i.e., quite a bit) during the observed class (M=2.68, SD=.84).
- Describing and explaining: Students reported that this cognitive activity occurred "some" of the time (M=2.03, SD=.93).
- Applying theories: Students reported that this cognitive activity occurred with relative frequency (M=2.62, SD=.91).
- Breaking ideas into constituent parts: Students reported that this cognitive activity occurred between "quite a bit" and "some" (M=2.42, SD=.90).
- Creating new ideas: Students reported that this cognitive activity occurred "some" of the time (M=2.16, SD=.93).
- Evaluate: Students reported that this cognitive activity occurred "some" of the time (M=2.19, SD=.97).
- Make connections with the real world: Students reported that this cognitive activity occurred between "quite a bit" and "some" (M=2.42, SD=.94).

Autonomous learning environment

Self-determination theory is an influential theory of motivation that focuses on the degree to which individuals are self-motivated and regulate their own behavior (e.g., Deci & Ryan, 1985). An important feature of motivation is feeling a certain degree of autonomy over one's behavior, and researchers have begun to explore the degree to which instructors provide an autonomous environment for student learning. In this study, we use a 6-item scale from Williams and Deci (1996) called the Learning Climate Questionnaire. Results for the entire study sample (n=248) indicate that students tend to agree with statements such as "My instructor encourages me to ask questions" (M=3.43, SD=.95) on a scale of 1-5.

Preliminary findings from student focus groups

In addition to these survey results, we also held focus groups with students in six courses. A preliminary review of the data indicates that students generally preferred teaching that was engaging and interactive, and they preferred class activities and material that were clearly linked to assessments. Additionally, students reported their in-class engagement in general terms such as "following along" or "being interested" instead of the more detailed categories of cognitive engagement as reported above.

RESULTS: Student study habits

In Table 5, descriptive data are provided about student study habits at the institution and disciplinary level. Descriptions of these items and scales are discussed in detail following.

TABLE 5: ***Results from survey of students' study habits and learning styles***

	Institution-Wide (13 courses)		Biological Sciences (6 courses)		Physics (2 courses)		Earth/Space Sciences (4 courses)		Mechanical Engineering (1 course)	
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
Learning style (1=never, 5=always)										
Surface strategy	256	3.02 (.70)	141	3.01 (.70)	70	3.02 (.71)	30	3.08 (.82)	15	3.01 (.51)
Deep strategy	256	3.00 (.68)	141	3.03 (.63)	70	3.01 (.74)	30	2.82 (.85)	15	3.04 (.47)
Perceived competence (1=not at all, 5=extremely)										
Perceived competence scale	247	3.85 (.87)	136	3.83 (.81)	68	3.92 (.98)	28	3.88 (.91)	15	3.78 (.78)
Study habits: Hours per week (1=less than 2 hrs, 2= 2-5 hrs, 3= 6-10 hrs, 4=11-15 hrs, 5= 16+ hrs)										
Hours per week: general	265	3.68 (1.19)	145	3.66 (1.16)	70	3.6 (1.19)	33	3.52 (1.37)	17	4.41 (.79)
(1=less than 1 hr, 2= 1-2 hrs, 3= 3-4 hrs, 4=5-6 hrs, 5= 7+ hrs)										
Hours per week: this course	265	2.41 (1.11)	145	2.26 (.98)	70	2.30 (.89)	33	2.18 (1.10)	17	4.65 (.60)
Study habits: Best practices (1=very little, 4=very much)			145							
Uses practice tests	256	2.58 (1.06)	141	2.62 (1.02)	70	2.54 (1.13)	30	2.33 (1.18)	15	2.80 (.86)
Uses practice tests over time	256	2.00 (1.06)	141	2.01 (1.03)	70	2.00 (1.06)	30	1.80 (1.00)	15	2.33 (1.35)
Uses different types of problems	256	2.80 (.99)	141	2.76 (.95)	70	2.99 (.96)	30	2.37 (1.16)	15	3.13 (.83)
Re-reads material	256	2.90 (1.01)	141	2.91 (.98)	70	2.97 (1.01)	30	2.77 (1.17)	15	2.67 (.90)
Study resources: General (1=very little, 4=very much)										
Other people	263	2.16 (.97)	145	2.00 (.90)	70	2.26 (1.02)	31	2.29 (1.16)	17	2.82 (.73)
Textbook	263	2.87 (1.09)	145	3.06 (.94)	70	3.07 (1.09)	31	1.87 (1.18)	17	2.24 (.90)
Notes taken in class	263	2.81 (1.02)	145	2.94 (.93)	70	2.76 (1.08)	31	2.16 (1.07)	17	3.06 (.97)
Lecture notes posted online	263	3.08 (.99)	145	3.21 (.90)	70	2.90 (1.11)	31	3.06 (1.06)	17	2.71 (.92)
Online course discussions	263	1.53 (.85)	145	1.52 (.82)	70	1.66 (.98)	31	1.29 (.69)	17	1.47 (.72)
Other online resources	263	2.14 (1.01)	145	1.99 (.91)	70	2.21 (1.10)	31	2.77 (1.12)	17	1.94 (.90)

Learning style

One of the more influential theories of student learning is that there exists particular "styles" of learning. This area of research emphasizes the students' own agency and decision-making in contrast to more cognitively-oriented theories, which some argue are de-contextualized accounts of learning that ignore students' beliefs and motives, the particular course being studied, and so on. Biggs, Kember and Leung (2001) refined the Study Process Questionnaire (SPQ), which measures both "deep" and "surface" approaches to learning, much like the aforementioned ASSIST survey. In the SPQ, items ask students to indicate how much they agree with statements such as "I find at times studying gives me a feeling of deep personal satisfaction" (deep learning style) and "My aim is to pass the course while doing as little work as possible" (surface learning style). For the study sample, the 256 students answering these questions did not exhibit much variation between the two types, with students reporting a surface strategy (M=3.03, SD=.70) and a deep strategy (M=3.00, SD=.68) in almost equal terms (on a scale of 1-5). These similarities held across disciplinary groups.

Perceived competence

Another feature of Deci and Ryan's (1985) self-determination theory is perceived competence. In educational settings, perceptions of one's own competence in a particular course is theorized to be important because it helps people attain their goals, persist in the face of challenges, and gives them a sense of satisfaction. For this study, we used the Perceived Competence Scale, which is a 4-item scale consisting of items such as "I feel confident in my ability to learn this material." The 247 students who answered this question reported feeling "somewhat" competent in relation to the target course (M=3.85, SD=.87) on a scale of 1-5.

Study habits: General

In the survey we also elicited the number of hours students spend studying for their entire course load and for the target course being currently studied. Across the sample, students on average spent 6-10 hours a week studying overall, and 1-2 hours on the target course. A notable exception is the mechanical engineering sample, where engineering students reported studying 5-6 hours a week on the target course. However, the students in this particular course were part of a single "course" that actually was comprised of multiple courses, such that their responses may be in reference to their study habits for more than a single course.

Study habits: Best practices

A recent review of the psychological literature on student study habits identified studying techniques that had strong evidence for supporting student learning (Dunlosky et al., 2013). In the student survey, we asked about the regularity with which they engaged in some of these "best practices" for studying in the target course, with the most frequently used study method being re-reading material (M=2.90, SD=1.01) followed by using different problem types (M=2.80, SD=.99), taking practice tests (M=2.58, SD=1.06), and taking practice tests that are spread out over time (M=2.00, SD=1.06) on scales of 1-4.

Study resources

One of the primary interests motivating this study was the type of resources that students draw upon when studying. This interest is largely based upon the increasing number of learning resources or tools that students have to draw upon, including in-class materials, out-of-class activities and assignments, course websites, and other formal online resources and social resources. For general resources, students in the sample reported that they most frequently used lecture notes that the instructor posted online (M=3.08, SD=.99), the textbook (M=2.87, SD=1.09), and notes taken in class (M=2.81, SD=1.02) on scales of 1-4.

Preliminary themes: Student focus groups

A preliminary review of the data from the focus groups in reference to student study habits revealed a few core themes. First, students generally are studying only in preparation for quizzes or exams, and often students cram during the few days prior to an assessment. Second, the lack of regular studying besides exam preparation is explained as inevitable given the course load that is carried by many students. Some felt that with the sheer amount of readings, problem sets, and exams they faced on a regular basis, there was no time for studying for the sake of studying. Third, in rare cases do students have a quiet, controlled study environment; rather, they often study in situations with many distractions, especially those of a digital nature (e.g., cell phones, laptops with Internet access). A closer examination of these data will yield interesting insights into how contemporary students engage in studying.

RESULTS: Relationships among planning, teaching, and student experiences for a sub-sample of courses

Finally, we report these distinct datasets (faculty planning, classroom teaching, and student experiences) across 13 courses as a way to illustrate the ISOP framework in action. While each individual set of data sheds important light on teaching and learning, it is only when viewed together that we begin to obtain a more accurate and multi-faceted account of the different aspects of this process. Of course, with data on actual student learning gains and/or processes related to a specific topic or class, this picture would be even more complete.

First, we examine the relationship among one of the key indicators of student learning – cognitive engagement. Table 6 includes data about the types of cognitive engagement faculty expect for students in the observed class, the actually observed cognitive engagement, and student reports of the types of cognitive engagement they experienced in the observed class.

TABLE 6: ***Compilation of datasets across individual instructors regarding expected, observed, and reported cognitive engagement***

Course #	Discipline	Course Enrollment	Faculty Interview- Expectation for Cognitive Engagement	Faculty Observation - Observed Cognitive Engagement: Proportion of 2-min Intervals				Student Survey - Reported Cognitive Engagement: Mean scores on 1(observed very little) to 4 (observed very much) scale			
				RRI	**PS**	**CR**	**CN**	**RRI**	**PS**	**CR**	**CN**
1	Physics	249	Struggle with material; talk with peers; make sense of problems	.21	.33	0	0	2.7 (.66) n=46	2.78 (.78)	2.2 (.91)	2.13 (.93)
2	Biology	218	Follow along, think about the material	.62	.37	0	.08	2.8 (.77) n=15	2.8 (.86)	2.27 (1.03)	2.33 (.61)
3	Geology	16	Actively thinking	.37	.23	0	.21	2.8 (.83) n=5	2.8 (1.09)	2.4 (1.34)	2.4 (.54)
4	Biology	210	Active questioning, active thinking	.88	0	0	.2	2.88 (.82) n=33	2.79 (.82)	1.94 (.93)	2.48 (.90)
5	Geology	55	Active thinking	.42	.20	0	.89	3.22 (.83) n=9	1.78 (1.3)	1.89 (.78)	2.33 (.86)
6	Biology	78	Critical thinking; applying knowledge	.47	.14	.04	.11	2.86 (.86) n=14	1.86 (.86)	2 (.55)	1.93 (.73)

Course #	Discipline	Course Enrollment	Faculty Interview- Expectation for Cognitive Engagement	Faculty Observation - Observed Cognitive Engagement: Proportion of 2-min Intervals				Student Survey - Reported Cognitive Engagement: Mean scores on 1(observed very little) to 4 (observed very much) scale			
				RRI	**PS**	**CR**	**CN**	**RRI**	**PS**	**CR**	**CN**
7	Mech Eng	165	Following along	.15	0	0	.34	2.47 (1.00) n=17	2.65 (.93)	2.35 (.99)	2.76 (.97)
8	Geology	49	Active questioning	0	.85	.42	0	1.8 (1.13) n=10	2.8 (1.02)	2.4 (.96)	2.2 (.91)
9	Biology	225	Come prepared and be actively thinking	.42	0	0	.31	2.64 (.69) n=33	2.7 (.84)	2.36 (.99)	2.52 (.90)
10	Geology	294	Critical thinking	.5	.35	0	.57	2.5 (.76) n=14	2 (.67)	1.86 (.86)	2.5 (.85)
11	Biology	197	Critical thinking and following along	.72	.10	0	.10	2.73 (.70) n=15	2.73 (.59)	2.13 (.99)	2.87 (.99)
12	Biology	112	N/A	.45	.2	0	.7	2.7 (.90) n=37	2.78 (.85)	2.11 (.89)	2.54 (1.01)
13	Physics	159	Integrating information	.45	.0	.0	.25	2.64 (.90) n=25	2.52 (1.08)	2.32 (1.06)	2.44 (1.12)

Note: Cognitive engagement codes for both classroom observations and student surveys: RRI (reciting and/or memorizing facts), PS (problem solving), CR (creating new ideas or products), and CN (making connections to the real-world). Note that student surveys did not include these codes named as such.

In this table one can read across individual rows to analyze the relationships among faculty expectations, their actual practice, and student experiences. For example, in a large enrollment physics course, the instructor expected students to struggle with the material, talk with their peers, and make sense of problems. The TDOP observation indicated that the types of cognitive engagement available to students were primarily receiving and memorizing information (in 21% of all two-minute intervals) and problem-solving (33%). After the observed class, 46 students reported that the types of cognitive experiences they most frequently observed in class were receiving and memorizing information (M=2.7, SD=.66) and problem solving (M=2.78, SD=.78). these data indicate an alignment between faculty expectations for student cognitive engagement regarding problem solving, but less of an anticipation for the amount of time they actually asked students to receive and memorize information.

Second, we compare data across a selection of observed faculty teaching practices, student perceptions of teaching quality, and student preferences for learning styles. This comparison is intended to explore the relationship between student perceptions of effective or high-quality teaching and specific teaching practices. Further, considering student preferences for learning (e.g., deep or surface) provides some important context to this comparison (see Table 7).

TABLE 7: ***Compilation of datasets across individual instructors regarding observed teaching, student perception of teaching quality, and student preferences for learning***

Course #	Discipline	Course Enrollment	Faculty Observation (Observed Teaching Practices: Proportion of 2-min Intervals)						Student Survey (Perceived Teaching Quality: Mean scores on 1-5 scale)			Student Survey (Reported Learning Styles: Mean scores on 1-5 scale)	
			LPV	**LHV**	**SGW**	**ORG**	**EMP**	**DQ**	**Clarity**	**Org**	**Challenge**	**Support Understanding (Deep)**	**Transmit Information (Surface)**
1	Physics	249	.57	.57	0	.12	.03	.60	3.08 (.90) n=48	3.89 (.87)	2.91 (.72)	2.96 (.72) n=43	3.95 (.70)
2	Biology	218	.91	.25	.20	.29	.16	.54	3.51 (.71) n=17	3.87 (.68)	2.72 (.58)	3.28 (.82) n=13	4.10 (.84)
3	Geology	16	0	.92	.17	.13	.05	.76	4.16 (.68) n=5	4.28 (.26)	4.23 (.81)	3.8 (.77) n=5	3.8 (.76)
4	Biology	210	.93	.17	.2	.17	.06	.37	3.65 (.76) n=35	4.38 (.54)	2.87 (.65)	3.16 (.72) n=32	3.60 (.78)
5	Geology	55	.98	0	0	.17	.04	.38	3.72 (.61) n=10	4 (.73)	2.69 (.79)	2.85 (.71)	3.80 (.97)
6	Biology	78	.78	.10	.15	.10	.05	.59	3.55 (.75) n=15	3.91 (.55)	2.86 (1.03)	3 (.80) n=13	4.08 (.89)
7	Mech Eng	165	.34	.65	0	.10	.04	.06	3.79 (.95) n=17	3.88 (1.08)	2.77 (.76)	3.15 (.56) n=15	3.78 (.67)
8	Geology	49	.06	.18	.89	.06	.02	.14	3.6 (.88) n=10	3.66 (1.16)	3.65 (.70)	2.9 (1.13) n=10	2.64 (1.18)
9	Biology	225	.97	0	.04	.10	.17	.70	3.90 (.64) n=55	4.08 (.55)	3.17 (.54)	3.07 (.60) n=29	3.53 (.82)
10	Geology	294	.78	0	.21	.07	0	.5	3.56 (.74) n=15	4.11 (.69)	2.73 (.78)	2.8 (.72) n=10	4.1 (.90)
11	Biology	197	.85	.12	.0	.19	.08	.17	4.27 (.71) n=15	4.54 (.52)	3.16 (.85)	2.67 (.80) n=15	3.8 (.96)
12	Biology	112	1	0	0	.15	0	.35	3.67 (.56) n=37	3.92 (.59)	3.38 (.70)	3.29 (.68) n=35	3.38 (.99)
13	Physics	159	.82	.20	.00	.14	.16	.11	4.43 (.48) n=28	4.61 (.39)	3.04 (.75)	3.23 (.78) n=25	3.70 (.96)

Note: Classroom observation codes include teaching methods, pedagogical moves, and student-teacher interactions. Teaching method codes include: LPV (lecturing with pre-made visuals), LHV (lecturing with hand-made visuals), and SGW (small group work). Pedagogical moves include: ORG (organizational markers) and EMP (emphasis). Student-teacher interactions include: DQ (display questions).

Again, for the large enrollment physics course, the instructor was observed lecturing over 50% of the time, using both pre-made visuals (e.g., PowerPoint) and hand-made visuals (e.g., board work). The instructor also utilized pedagogical strategies such as using organizational markers, such as verbally declaring when she was moving to a new topic, as well as the frequent (in 60% of all two-minute intervals) posing of display questions (i.e., seeking new information) to students. This instructor then can be viewed as heavily relying on lecturing, but with frequent pausing for questions and the intermittent "marking" of transitions between topics. 48 students perceived the instructor as having a relatively high degree of organization (M=3.89, SD=.87). This group of students also reported learning preferences that favored a surface approach (i.e., transmitting information). While these data do not provide indications of causal relationships among faculty interview, faculty classroom observations, and student survey datasets, they do provide interesting insights into the teaching and learning dynamic.

Conclusions

Based on the data provided in this report, several conclusions can be drawn about the use of data in curricular decision-making. First, while formal systems exist for creating new courses and introductory-level service courses, the curriculum design process is largely ad-hoc and dependent upon individual instructors. Second, while teachers are regularly evaluated through end-of-term evaluations, the course itself and whether or not it has met pre-specified learning goals or even satisfied basic criteria, is rarely assessed in a systematic manner. However, some instructors and teams of instructors hold regular meetings or end-of-term gatherings to determine if the course met its goals. Third, a variety of data types are used to inform course planning both formatively and summatively. However, the rigorous tracking of data as part of a continuous improvement system is very rare, such that it is difficult to ascertain with any certainty if a course was successful. Further limiting the ability of administrators and faculty to make this determination is the sporadic articulation of measurable learning goals for students, without which there are no clear benchmarks upon which to assess course efficacy.

Policy Recommendations

In light of the growing efforts to improve the quality (and accountability) of undergraduate education via accreditation agencies, institutional mandates, or educational policies, it is not unlikely that at some point in the future faculty will be required to articulate and track measurable learning outcomes for students. Indeed, this is already the case in engineering fields, where the Canadian Engineering Accreditation Board has recently developed new metrics for ensuring program quality that include measurable learning goals. We applaud the push for encouraging educators to reflect on formative and summative data towards the ultimate goal of improving teaching and learning, and recognize the important role that data-driven decision making can play in such improvement. Yet we also share the trepidation of some that these systems can become onerous, top-down mechanisms for estimating "quality" that miss the mark.

In ideal cases, the development of these accountability systems and metrics for tracking educational quality will be the product of a collaborative exercise where faculty, students, administrators, and policymakers collectively determine appropriate systems. However, in many cases, such as the implementation of accountability systems in the U.S. in response to the NCLB legislation, these systems are developed by administrators with little input or buy-in from classroom teachers. Teachers and administrators in postsecondary settings are generally at an advantage in that no legislation yet exists on the order of NCLB, such that institutions and departments are largely left to their own devices in regard to assessment and evaluation. We recommend that teachers and administrators take advantage of this situation and "get ahead of the curve" by developing their own data systems. We view this as an ideal situation because data systems are most successful when the people who are most knowledgeable of the courses and students (i.e., the faculty) are involved in designing these systems. and they are more likely to be successful.

1. Establish mechanisms for collecting, analyzing, and reflecting upon course data

First, we recommend that each department institute some type of continuous improvement system and/or feedback loop where instructors and administrators have the means to continually improve their own courses and teaching skills, and they have the ability to track progress towards student learning goals. Key components of such systems include: (a) the articulation of measurable learning goals, (b) appropriate assessments and/or types of data to track progress towards these goals, and (c) systems for storing and managing these data. Local leaders will have to determine whether or not these systems should operate at the department, course sequence, or individual course levels, and also whether or not compliance with these systems is voluntary or mandated. If local leaders choose the latter, which would increase the utility and benefits associated with data-driven decision making, the systems should be carefully designed in collaboration with the affected faculty.

Further, articulating learning goals should be done initially by faculty who are directly involved in teaching a course. Identifying appropriate goals is a difficult enterprise, particularly if they pertain to higher-order cognitive skills that are challenging to measure, and faculty familiar with a course are in a good position to determine which goals are most appropriate for a particular set of topics.

2. Ensure that these systems are not punitive or evaluative, but instead are designed to support teacher growth and development.

Ultimately, the successful institution of data systems in postsecondary classrooms will require a clear statement by administrators that the systems are not intended as assessments of instructor quality per se, but instead as means to help instructors monitor their courses in a rigorous fashion and benefit from the analysis of these data to grow and develop their pedagogical skills and the learning skills of their students. It is advisable to respect the autonomy and disciplinary expertise of faculty, and to not adopt an overly bureaucratic stance by mandating these systems using a top-down approach to governance. Instead, in order for these systems to truly facilitate reflective practice that leads to quality instruction, faculty must view them as aids to their practice and not as a means for evaluation or punishment. That said, it is not unlikely that at some point colleges and universities will be asked to evaluate the quality of teaching within their institutions, and if this comes to pass, all faculty will need to "get on board" with assessment. Ideally, faculty themselves will be involved in the articulation of these assessment systems, rather than the situation that has transpired in K-12 schools were classroom teachers generally have little input into how they are evaluated for promotion and retention.

3. Identify and support an individual to help faculty interpret and apply these data to course improvement

Finally, the creation of these goals and systems should not put more work on the instructor, who is already very busy, requiring the designation of support staff (preferably within departments or disciplines) who can assist faculty with the collection and interpretation of these data. Mandinach (2012) calls these support staff "data helpers," and they should have a high degree of literacy with pedagogical data. These staff persons can help to design formative and summative feedback systems for individual courses, which are quite difficult to establish and operate in an effective manner. This person can work in close consultation with the lead instructor who best knows their course, the material, and the students. Then, once data are collected, this person can help faculty to interpret these data and make them actionable.

Next Steps

Data reports are currently being provided to administrators and faculty at Institution A and two other participating institutions using data from the spring of 2013 (Wave 1). The data are being provided to two sets of individuals: (1) administrators and faculty engaged in curriculum design at the department and/or program level, and (2) individual instructors who participated in this study. The next steps for the TPDM study are to conduct follow-up interviews with individuals receiving these reports to ascertain whether or not they were useful, enhanced local decision-making practices, and to obtain any critiques or feedbacks regarding our approach. In addition, we will also be training interested individuals to use the TDOP website and other components of the ISOP framework. Then, in the spring of 2015, we will administer the ISOP framework at these same three institutions (Wave 2) in order to track any changes in course planning, teaching, and student experiences—and especially, the use of data in curricular decision-making.

References

Arum, R., & Roksa, J. (2010). *Academically adrift: Limited learning on college campuses.* University of Chicago Press.

Biggs, J., Kember, D., & Leung, D.Y.P. (2001). The revised two-factor Study Process Questionnaire: R-SPQ-2F. *British Journal of Educational Psychology, 71,* 133-149.

Bill & Melinda Gates Foundation (2012). *Asking Students about Teaching: Student Perception Surveys and Their Implementation.* Seattle, WA: Bill & Melinda Gates Foundation. Retrieved November 13, 2012, from www.metproject.org/downloads/Asking_Students_Practitioner_Brief.pdf.

Bok, D. C. (2008). *Our underachieving colleges: A candid look at how much students learn and why they should be learning more.* Princeton University Press.

Camburn, E. M. (2012). Review of asking students about teaching. Boulder, CO: National Education Policy Center. Retrieved April 4, 2013 from http://nepc.colorado.edu/thinktank/review-asking-students

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York: Plenum Press

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest, 14*(1), 4-58.

Entwistle, N., & Tait, H. (1990). Approaches to learning, evaluations of teaching, and preferences for contrasting academic environments. *Higher education, 19*(2), 169-194.

Guarino, C. & Stacy, B. (2012). *Review of "Gathering Feedback for Teaching."* Boulder, CO: National Education Policy Center. Retrieved October 28, 2012 from nepc.colorado.edu/thinktank/review-gathering-feedback.

Halverson, R. (2003). Systems of practice: how leaders use artifacts to create professional community in schools. *Educational Policy Analysis Archives, 11* (37), 1-35.

Hora, M. T., & Ferrare, J. J. (2012). Instructional Systems of Practice: A Multidimensional Analysis of Math and Science Undergraduate Course Planning and Classroom Teaching. *Journal of the Learning Sciences, 22* (2), 212-257.

Mandinach, E. B. (2012). A perfect time for data use: Using data-driven decision making to inform practice. *Educational Psychologist, 47*(2), 71-85.

Murray, H.G. (1999). Low-inference teaching behaviors and college teaching effectiveness: Recent developments and controversies. In J.C. Smart (Ed.), Higher education: Handbook of theory and research. Volume 15. New York: Agathon.

Perry, R. P. & Smart, J. C. (Eds.). (1997). *Effective teaching in higher education: Research and practice.* New York, NY: Agathon.

Saroyan, A., & Snell, L. S. (1997). Variations in lecturing styles. *Higher Education, 33*(1), 85-104.

Tait, H., Entwistle, N. J., & McCune, V. (1998). *ASSIST: a reconceptualisation of the Approaches to Studying Inventory.* In C. Rust (ed.) *Improving students as learners.* Oxford: Oxford Brookes University, The Oxford Centre for Staff and Learning Development.

Umbach, P. D. (2004). Web surveys: Best practices. *New Directions for Institutional Research, 2004*(121), 23-38.

Wabash National Study of Liberal Arts Education. Student Experiences Survey. http://www.liberalarts.wabash.edu/study-overview/

Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychological values by medical students: A test of self-determination theory. *Journal of Personality and Social Psychology, 70,* 767–779.

Resources

For more information about the TPDM study see (tpdm.wceruw.org) and for information about the TDOP see (tdop.wcer.uw.org).

For more information about the ASSIST instrument and related research on student learning see (www.etl.tla.ed.ac.uk/project.html).

Acknowledgments

This research was supported by the National Science Foundation (Award #1224624). Authors are solely responsible for publication content. Any opinions, findings, and conclusions or recommendations expressed in this material do not necessarily reflect the views of the National Science Foundation or representatives of the University of Wisconsin–Madison.

Appendix A

Description of sample for interviews and observations

	Interview n	Percentage %
TOTAL	19	100%
Sex		
Female	8	42%
Male	11	58%
Discipline		
Mechanical Engineering	3	16%
Physics	3	16%
Biology	9	47%
Geoscience	4	21%
Level of course		
Lower division	14	74%
Upper division	5	26%
Size of Course		
50 or less	2	10%
51-100	2	10%
101-200	7	37%
201 or more	8	42%
Position type		
Lecturer/Instructor	11	58%
Assistant Professor	3	16%
Associate Professor	2	10%
Professor	3	16%

Appendix B

Description of TDOP inter-rater reliability scores for analysts

	Teaching Methods	Pedagogical Moves	Interactions	Cognitive Engagement	Instruct. Techn.
Analyst 1/Analyst 2	.90	.85	.83	.74	.94
Analyst 1/Analyst 3	.82	.81	.73	.78	.90
Analyst 1/Analyst 4	.89	.74	.79	.71	.90
Analyst 2/Analyst 3	.83	.80	.81	.75	.89
Analyst 2/Analyst 4	.84	.75	.79	.77	.89
Analyst 3/Analyst 4	.80	.73	.72	.74	.91

Note: Figures represent Cohen's kappa scores averaged across two observed class periods (one in biology, one in physics). Cohen's kappa s is an index of inter-rater reliability measures the level of agreement between two sets of dichotomous ratings, while taking into account the possibility that agreement can take place by chance.

Appendix C

Description of student survey alpha scores

Survey scales (# of individual items)	Responses	Cronbach's Alpha
Perceptions of teaching quality		
Clarity (5)	282	.848
Organization (5)	282	.873
Challenge (6)	282	.774
Preferences for teaching		
Supporting understanding/deep (4)	248	.623
Transmitting information/surface (3)	248	.696
Autonomous learning environment (6)	248	.917
Perceived competence (4)	247	.916
Learning style		
Supporting understanding/deep (5)	256	.557
Transmitting information/surface (5)	256	.552

WISCONSIN CENTER FOR EDUCATION RESEARCH

UNIVERSITY OF WISCONSIN-MADISON
1025 West Johnson Street, Suite 785 | Madison, Wisconsin 53706

Tracking the processes of data-driven decision making in higher education (TPDM)
tpdm.wceruw.org | If you have any questions about the TPDM project, please contact Matthew Hora at hora@wisc.edu.

This material is based upon work supported by the National Science Foundation under Grant No. DUE-1224624. 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 National Science Foundation.