

**A Visual Approach to Helping Instructors Integrate, Document, and Refine Active Learning:
Supplemental Materials**

These *Supplemental Materials* include Part I: Examples of Suggested Instructional Changes Resulting from Discussion of OPAL Timelines and Part II: Development of the Observation Protocol for Active Learning (OPAL). Part II contains information on the development of OPAL and OPAL timelines, the training OPAL observers, and a full listing of all OPAL codes.

Part I. Examples of Suggested Instructional Changes Resulting from Discussion of OPAL Timelines

In the current study, we employed the OPAL timeline in peer review of teaching that occurs during Teaching Center consultations with a requesting instructor. The reviewers are faculty colleagues who are members of The Teaching Center staff and experienced instructors. These consultations include discussion of i) the instructor’s objectives for the course, ii) segments of a video-recorded class session that has been reviewed by the Teaching Center staff member, iii) an OPAL timeline from the same class session, and iv) the reviewer’s suggestions for potential instructional changes that the instructor could implement to meet his or her stated objectives (*citation removed for blinded review*). Table 1 includes five sample cases that illustrate how patterns observed in an OPAL timeline prompted the reviewer to suggest the noted potential instructional changes. These cases are provided as examples that illustrate how review of an OPAL timeline can lead to potential instructional changes, during either peer review or self-review of teaching.

Table 1. *Sample cases of peer-review consultation discussions.*

Cases	Instructor’s Objective	Observed in OPAL Timeline	Interpretation	Potential Instructional Change (Suggested by Reviewer)
1	Increase amount of interaction—between instructor and students, and among students.	Instructor asked several questions, but with limited response from students.	Instructor may not have had a plan for prompting students to answer questions in discussion with one another and for	Implement an interactive lecture style that includes, at regular intervals, questions that students will discuss in small groups or pairs.

			connecting student responses to the concepts in the lecture.	Include specific strategies for eliciting student responses, then connecting these responses to the next idea or step in the lecture.
2.	Transition from “traditional” lecture (without interaction) to interactive lecture.	Instructor asked one question during lengthy lecture segment (30 minutes of continuous lecturing).	Instructor may not yet be comfortable with an interactive lecture style and as a result fell back on “traditional” lecture as the default mode.	Integrate more frequent questions and opportunities for students to engage, preferably from the start of the class session, so that instructor and students become comfortable moving back and forth between lecture and interaction.
3.	Integrate demonstrations of experiments to help students develop understanding of key scientific concepts.	Instructor included a video showing a brief experiment, but did not provide set-up or follow-up questions to prompt students to respond to or discuss the demo.	Students may not have understood the concepts that the instructor intended the video to reveal or clarify.	Ask students pre- and post-demo questions, such as “ <i>Can you predict what will happen?</i> ” (pre), “ <i>What happened that you did not expect?</i> ” (post), and “ <i>Why did this result occur?</i> ” (post).
4.	Integrate group work to help students learn from one another.	Instructor gave students a problem to work on in groups, but did not walk around the room to observe groups during the group work.	Instructor may not have a clear plan for facilitating group work without interrupting students’ collaboration with one another.	Circulate around the room to observe group work and help students stay focused and to answer any questions about instructions for group work, but do not intervene into group discussions.
5.	Integrate clicker questions into lectures to help students explore complex concepts and solve problems.	Instructor moved from discussion of clicker question back to lecture without summarizing important points from the discussion; students did not take notes.	Students may not have identified and recorded main points from discussion of clicker question and how these points were related to the lecture content.	Integrate brief summaries of key points from discussion, using the chalkboard to record these points and to communicate to students that they should include these points in their notes.

Part II: Development of the Observation Protocol for Active Learning (OPAL)

As described in the article, the OPAL tool includes Behavior Codes and Activity Levels (note-taking and attention), resulting in nominal and ratio-level data. While we recorded both nominal and ratio-level data in the current study, it should be noted that this complexity can be removed and only nominal-level data can be recorded. Either method will yield similar results and a similar timeline.

The behavior codes and activity levels of note-taking and attention were adapted from the Teaching Dimensions Observation Protocol (TDOP; Hora, 2013) and the Classroom Observation Protocol for Undergraduate STEM (COPUS; Smith, Jones, Gilbert, & Wieman, 2013), or developed by our research team during the pilot phase of this project. Both TDOP and COPUS use occurrence sampling to record the activities occurring in the classroom. COPUS also uses an “engagement” code, from which we derived the OPAL Activity Levels for note-taking and attention (Smith et al., 2013). Because it is difficult to quickly and accurately assess the percentage of students who are paying attention and taking notes, we adopted the method used by Smith and colleagues (2013), documenting Activity levels with recording intervals of <20%, 20-80% and >80% of students paying attention and taking notes. In addition, we added a “zero” level for note-taking. These respective ranges, while uneven, designate two extremes and a broad middle range that are readily distinguishable from one another, allowing the observer to indicate whether none of the students, very few of the students (<20%), a moderate amount of students (20-80%) or nearly all of the students (>80%) are taking notes or paying attention. We also use scan sampling to improve the consistency of the assessment of Activity Levels. We have found that more consistent data are collected when the classroom is scanned at a designated point in time—in this case at the end of each two-minute interval. In addition, we assume that any typing or writing during class is note-taking, while acknowledging that this assumption necessarily involves some error.

In this study, we completed 144 observations using OPAL, between January 2014 and April 2015. Descriptive characteristics of the class sessions observed can be found in Table 2. The vast majority of the observations were conducted in person, during a “live” class session, but 14 were coded by watching video-recorded classes. OPAL can be used with both live and video-recorded observations, provided that the recording captures video of both the instructor and the students. In this study, the video-recorded classes were recorded for a purpose other than the current research and as a result focused almost exclusively on the instructor. Therefore, the OPAL data collected for these classes are less comprehensive than the data recorded for classes that were coded through in-person observation.

Table 2. *Characteristics of class sessions observed using OPAL, January 2015-April 2014.*

Course characteristics	Number of Observations (N)	% of Total Observations (Total Number of Observations = 144)
Course Enrollment		
< 75 students enrolled (large)	112	77.8
> 75 students enrolled (small)	32	22.2
Course Level		
Lower-level (100- and 200-level)	130	90.3
Upper-level (300- and 400-level)	14	9.7

Course Discipline		
STEM discipline: Science <ul style="list-style-type: none"> • Biology ($n = 18, 14.8\%$) • Chemistry ($n = 48, 39.3\%$) • Physics ($n = 47, 38.5\%$) • Psychology ($n = 9, 7.4\%$) 	122	84.7
STEM discipline: Engineering <ul style="list-style-type: none"> • Chemical Engineering ($n = 6, 66.7\%$) • Computer Science ($n = 3, 33.3\%$) 	9	6.3
Non-STEM discipline <ul style="list-style-type: none"> • Economics ($n = 6, 46.1\%$) • English ($n = 1, 7.7\%$) • General Studies ($n = 2, 15.4\%$) • Interdisciplinary Studies ($n = 3, 23.1\%$) • Political Science ($n = 1, 7.7\%$) 	13	9.0

Note. Within each main category (Course Enrollment, Course Level, and Course Discipline), values in the

N column (number of observations) add up to 144 and values in the percent column add up to 100.

Lower-level courses are those with primarily first- and second-year students enrolled, and upper-level courses are those with primarily third- and fourth-year students enrolled.

OPAL Timeline

As described in the article, we developed—for instructors conducting peer review and self-review of teaching—a streamlined version of our OPAL timeline that simplifies the codes and includes brackets labelled to identify recognizable segments of the class session. While a sample streamlined timeline showing the first 30 minutes of an observed class is included in the article, a complete detailed timeline and a complete streamlined timeline (all recorded for the same class session) can be found in Figures 1 and 2, below. The streamlined timeline differs from the detailed version in two ways. First,

similar codes are consolidated by the research team. For example, OPAL codes for questions, such as Pose Question Verbally (PQv) and Brief Question to Check Comprehension (ChQ), are consolidated and marked on the streamlined timeline as Question (Q). Second, we add brackets and labels to indicate recognizable segments of pedagogical strategies, e.g. Lecture with Questions, Lecture, Interactive Lecture or Problem Solving with Group Work. These brackets and labels are added for the purpose of visual aids for the instructor; they are not recorded in the OPAL data that are analyzed for research purposes. This step is performed by a senior research-team member, who examines the detailed timeline and identifies segments of related instructor behaviors. In cases where it is not immediately clear how to label a segment, multiple members of the research team discuss possibilities and determine the appropriate label. In both the detailed and the streamlined version, the ratio-level data (number of questions and answers) are shown on a separate timeline (denoted as Number of Q's and A's). In this timeline, each question or answer is represented by one cell, and questions and answers by students and instructors are merged into a single number for each two-minute interval. Tables 3 and 4 below provide the list of streamlined codes used in the figures and examples of bracket labels, respectively. The list of streamlined codes in Table 3 references the original OPAL behavior codes from which each streamlined code is derived.

Table 3. List of all OPAL codes matched with streamlined OPAL codes.

OPAL Codes		Corresponding Streamlined Codes	
Instructor Codes	Student Codes	Code	Description
AnQ	AnQ, GrA, Pan	A	Answering
AA		AA	Acting animation
Ad _c , Ad _T , Ad _A		Ad	Administration
ADV, PDV		D	Demonstration/video
FD	WDI, WDp	Dc	Discussion
Sfu		F	Follow-up
Dfu		FQ	Follow-up with questions
	QG, WG	G	Group-work
Lint		IL	Interactive lecturing
Lec,		L	Lecturing
LHV		LH	Lecturing with hand-made visuals
LPV		LP	Lecturing with pre-made visuals
	L _i , Lad, L _s	Li	Listening
MG _T	MG _S	M	Moving around room
O	O	O	Other
PSw, PSb, PSv	Ind, WR, CW	PS	Problem solving
ChQ, PQv, PQb	SQ	Q	Questioning
	R	R	Responsiveness
	TQ, SP	TP	Test or presentation
	V _H , V _T	V	Voting
	W	W	Waiting

Table 4. Sample streamlined timeline bracket labels, with OPAL codes typically included in each bracketed segment.

Bracket label	Typical codes in this bracket
Admin/Administration	Ad _C , LAd
Lecture	Lpv, LHV, L _I
Lecture with questions	Lec, PQv, SQ, AnQ, V _T
Problem solving	PSb, PSv, Sfu, Dfu, Ind
Problem solving with groupwork	PSb, PSv, Sfu, Dfu, QG, WG
Demonstration	PDV, L _I
Active demonstration	ADV, CW
Video	PDV, L _I
Interactive Lecture	Lint, L _I
Test/Quiz	Ad _T , TQ
Student presentations	FP, SP
Discussion	FD, L _S , WDI, WDP

Note. All code descriptions can be found in the full list of OPAL codes below.

Training of OPAL Observers

We use a systematic approach to training our observers. While extensive knowledge of teaching and pedagogy is not needed to successfully code, all of our observers are Ph.D. level staff, postdoctoral fellows, or faculty in a science, technology, engineering, or mathematics (STEM) discipline and have an interest in teaching. These observers' knowledge of the STEM education observation literature varies, with some having relatively high levels of knowledge and others being familiar only with OPAL codes and procedures. The training occurs over multiple days, and lasts 5-8 hours in total. As seen in Table 5 below, we use a combination of videos and live observations in our training, and each coding session is followed by a discussion of the session with a previously trained observer. We use videos at the beginning of the training process to allow the trainee to pause, stop, and replay portions of the video as needed.

Moreover, we begin with a video of a class in which lecture is the primary method—a format that tends to require the use of fewer codes than a class that incorporates multiple segments of active learning. Therefore, the trainees begin with a class that is relatively easier to code than classes they will code during later phases of the training and during data collection.

Table 5. *Template for training OPAL observers.*

Time	Suggested activity
1 hour	1. Trainees meet with trained observers to discuss and clarify codes. During meeting, trainees practice coding only the instructor using short video clips from the web.
50 minutes	2. On their own time, trainees code a 50-minute video of a traditional lecture-based course.
Up to 1 hour	3. Trainees meet with trained observers to discuss and compare coding of the video.
50 minutes	4. On their own time, trainees code a 50-minute video of a lecture-based course that incorporates active learning.
Up to 1 hour	5. Trainees meet with trained observers to discuss and compare coding of the video.
50 minutes	6. Trainees accompany trained observers to perform an in-person observation and independent coding of the observed class.
Up to 30 minutes	7. Trainees and trained observers discuss and compare respective coding of the observation.
50 minutes	8. Trainees accompany trained observers to perform a second in-person observation and independent coding of the observed class (if needed).
Up to 30 minutes	9. Trainees and trained observers discuss and compare respective coding of the observation (if needed).
As needed	10. Trainees and trained observers maintain communication throughout the semester.

Full List of OPAL Behavior Codes and Activity Levels

OBSERVATION PROTOCOL FOR ACTIVE LEARNING (OPAL)

BEHAVIOR CODES

Instructor Behaviors

Course Administration Codes

- **Ad_c**—General course administration (the instructor is giving students important reminders, dates, announcements, taking attendance, setting up A/V, returning work, etc.).
- **Ad_t**—The instructor is doing Test/Quiz administration for the class.
- **FP**—The instructor is facilitating student presentations.

Content Delivery Codes

- **Lec**--Lecturing; talking to the students and not using any visuals or demonstration equipment. Lec could also be marked if instructor has stopped using pre-made or handwritten visuals for more than one minute while still lecturing.
- **Lpv**--Lecture with pre-made visuals; talking to students while using visual aids that were made prior to the start of class (e.g. slides; pre-written chalkboard notes, etc.). Mark only when visuals are directly *referenced*; not when visuals are up but are not being used.
- **LHV**--Lecture with handwritten visuals. Instructor is talking to students while actively writing on board/projector/etc. or referencing previously written visuals from earlier in the same class session.
- **Lint**--Interactive lecturing; talking to students while asking multiple, interrelated questions in order to examine a complex problem, idea, issue, figure, or experiment (to name a few examples). Instructor integrates student answers into the lecture. Might take a few minutes to understand that Lint is occurring rather than PQ→AnQ, or be coded retroactively. Count number of questions and answers in the PQv and student AnQ codes.

Pedagogical Style Codes

- **MG_t**—Instructor is physically moving through class while lecturing or during an activity, guiding ongoing student work or asking/answering questions.
- **AA**—Acting/Animation. Instructor acting out a concept using their body.
- **ChQ***—Comprehension question. Ex. “Does anyone have questions?” or “Does that make sense?” (count number of times).
- **AnQ***—Listening to and answering student questions with entire class listening (count number of times).

Teaching Activities Codes

- **Ad_A**—Administration of activity or task (e.g. giving instructions, time left to complete activity, etc.).
- **FD**—Facilitate discussion. Instructor facilitating a whole class discussion, where whole class is engaged (i.e. students not discussing in small groups).
- **PQv***—Posing question to whole class regarding course content verbally (non-rhetorical). PQv’s tend to be quick interactions where no explicit time is given to students to think, discuss, or problem solve. Student codes frequently marked in response to PQ include AnQ, GrA, V_H, or V_T. (count number of times).
- **PQb***—Posing question to whole class regarding course content by displaying that question on the board, on PowerPoint, etc. (non-rhetorical). PQb’s tend to be quick interactions where no explicit

time is given to students to think, discuss, or problem solve. Student codes frequently marked in response to PQ include AnQ, GrA, V_H, or V_T (count number of times).

- **PSw**—The instructor presents a problem solving activity/task presented on a worksheet for students to complete. This could be a writing assignment, drawing assignment, problem solving activity, etc. PSw's tend to be longer than a simple PQ; students are given time to think, discuss, or problem solve before arriving at their answer. Student codes frequently marked in response to PS include Ind, WR, QG, or WG.
- **PSb**—The instructor presents a problem solving activity/task presented on board or PowerPoint for students to complete. PSb's tend to be longer than a simple PQ; students are given time to think, discuss, or problem solve before arriving at their answer. Student codes frequently marked in response to PS include Ind, WR, QG, or WG.
- **PSv**—The instructor presents a problem solving activity/task presented verbally for students to complete. PSv's tend to be longer than a simple PQ; students are given time to think, discuss, or problem solve before arriving at their answer. Student codes frequently marked in response to PS include Ind, WR, QG, or WG.
- **ADV**—Active demo/movie/video/simulation. Showing or conducting a demo, experiment, simulation, video, or animation while students are also actively engaged in the activity (interacting with each other or instructor, using equipment, etc.). Also mark any time prior to demo in which instructor is explaining demo/setting it up.
- **PDV**—Passive demo/ movie/video/simulation. Showing or conducting a demo, experiment, simulation, video, or animation with little to no interaction with students. Also mark any time prior to demo in which instructor is explaining demo/setting it up.

Activity Follow-up Codes

- **Sfu**—Summary follow-up. When instructor gives follow-up to an activity that is only a summary/explanation (i.e. little to no student interaction). May be difficult to tell if Sfu is occurring or if lecture has begun. Mark Sfu only when the activity is being directly referenced and no new material is being presented.
- **Dfu**—Discussion follow-up. When instructor gives follow-up to an activity by interacting with students (often asking them how they arrived at their answer). Mark each PQv, AnQ, etc. that occurs during Dfu.

Other

- **O**—Other. Explain in comments.

Student Behaviors

Non-Content or Non-Activity Codes

- **W**—Waiting (e.g. fixing AV problems, instructor otherwise occupied, etc.).

Student Assessment Codes

- **TQ**—Student taking a test or quiz.
- **SP**—Formal presentation by student(s).

Student Listening and Responsiveness Codes

- **L_r**—Listening to instructor speaking to whole class, while possibly taking notes (NOT when listening to instructor giving course administration or instructions to an activity).
- **LAd**—Listening to course administration or instructions to an activity given by instructor
- **L_s**—Listening to other students speaking, while possibly taking notes, etc. L_s should only be coded when the student response is lengthy (i.e. many sentences, longer than 1 minute).
- **R**—Responsiveness to instructor (e.g. laughing, collective sigh, etc.).

Student asks a question

- **SQ***—Student asks question (count number of times).

Student Responding Codes:

- **AnQ***—Student individually (and verbally) answers a question posed by the instructor with rest of class listening (count number of times).
- **GrA***—Students give group answer when the instructor poses a question to the class; the answer is given as a collective response rather than one individual student (count number of times).
- **Pan***—Student presents their answers to a problem or activity in front of the class via the board/large sticky note.
- **V_H***—Student votes for an answer by a show of hand (count number of times).
- **V_T***—Student vote/poll an answer by using the clicker or a phone app to register their answers (count number of times).
- **MG_s**—The students are moving around the classroom or physically moving their body or parts of their body in response to a prompt from the instructor.

Student prompted to think or work on activity/question/simulation/etc:

- **Ind**—Individual thinking/problem solving. Only mark when working on an activity/question/problem individually (i.e. not when only listening to lecture).
- **WR**—Student writing on paper or computer/tablet.
- **QG**—Discuss question or activity in groups. Ind could also be marked concurrently if more than half the class is not working in groups as instructed and are instead working individually.
- **WG**—working in groups on a problem or activity on a worksheet or handout.
- **WDI**—Whole class discussion instructor led. Class is engaged in whole class discussion/debate with students and instructor offering explanations, opinion, judgment, etc. to whole class, largely facilitated by instructor (This is typically a longer class activity more than 2 minutes).
- **WDP**—Whole class discussion with peer-to-peer interaction. Class is engaged in whole class discussion/debate with students and instructor offering explanations, opinion, judgment, etc. to whole class, with peer to peer interaction. Only code when it is clear that students are interacting with other students while engaged in a whole class discussion, otherwise code WDI. (This is typically a longer class activity more than 2 minutes).
- **CW**—Students are working with a physical model or on a non-word processing computer software program. This is an activity you might see in a studio classroom, active learning classroom, or mathematics or physics lab.

Other

- **O**—Other. Explain in comments.

ACTIVITY LEVELSAttention

- Mark the percentage of students paying attention to classroom instruction, or engaged in classroom activity. Mark as high (more than 80% clearly engaged), medium (between 20-80% clearly engaged), and low (less than 20% clearly engaged). Classroom should be scanned sampled near the end of each 2 minute segment to determine engagement level.

Note-taking

- Mark the percentage of students engaged in note-taking during classroom instruction. Mark as high (more than 80% clearly taking notes), medium (between 20-80% clearly taking notes), low (less than 20% clearly taking notes), or zero (no students taking notes). Classroom should be scanned sampled near the end of each 2 minute segment to determine note-taking level.

Note. All behavior codes are recorded as nominal-level data (i.e. either occurred or did not occur in the two-minute segment). In addition, the codes marked with an *are also recorded as ratio-level data in the current study (i.e. number of occurrences per two-minute segment are recorded).