Fostering self-reflection on library instruction: Testing a peer observation instrument focused on questioning strategies

http://dx.doi.org/10.11645/18.1.54

Eric Silberberg
Assistant Professor, Librarian for Instructional Design and Education, Queens College, City University of New York. Email: eric.silberberg@qc.cuny.edu. ORCID: 0000-0002-2216-7632.

Abstract

This study demonstrates that a library instruction observation instrument can effectively foster critical self-reflection among academic library faculty and staff on their teaching practices. The paper outlines the instrument's design, which gathers low inference observations on instructors' use of questioning as a pedagogical strategy based on recommendations from the LIS and education literature. To test and refine the instrument's design, the instructors' utilised the instrument to collect data from classes taught by five participating instructors, who, during post-observation interviews, engaged in thoughtful reflections on their class planning, student participation, and teaching philosophy. They also provided valuable critiques of the usefulness of the instrument. Through analyzing the observee reflections and the data from the observation instrument, this study aims to provide academic libraries with a method to incorporate an observation instrument in a peer observation program.

Keywords
critical information literacy; information literacy; information literacy instruction; pedagogy; reflective practice; research methods; teaching methods

1. Introduction

Academic librarians, who count teaching information literacy (IL) as a significant part of their responsibilities, often lack formal teacher training (Brecher & Klipfel 2014; Wheeler & McKinney,
While some may have had exposure to pedagogical concepts through graduate coursework, mastering effective teaching can only be achieved through practical experience. Therefore, opportunities for self-reflection play a crucial role in facilitating librarians’ professional growth as teachers (Houtman, 2019; Whitver, 2019).

One promising approach to facilitate feedback and self-reflection is peer observation (O’Leary, 2020), which academic libraries have adopted due to its ability to foster knowledge exchange, provide professional support, and cultivate reflective teaching practices (Alabi & Weare, 2014). However, observation is limited by the observer’s communication skills and potential biases. Moreover, observations often serve evaluative purposes associated with the reappointment or tenure processes, rather than being an integral part of a comprehensive professional development program (Kohut et al., 2007).

This present study addresses the following research questions: (1) How does the use of an observation instrument in peer observation facilitate critical self-reflection among library instructors on their teaching practices? (2) How does using low-inference questions during a peer observation facilitate self-reflection conversations with library instructors? To answer these questions, the study developed an observation instrument that collects low inference observations on how instructors use questioning in the classroom. Questioning is a fundamental instructional strategy that lies at the heart of effective teaching and learning. It serves as a powerful tool to engage students, elicit responses, and promote critical thinking (Cruickshank et al., 2006; Fusco, 2012; Jacobson & Xu, 2004; Song, 2019; Walsh & Sattes, 2015). This study argues that by focusing the observation instrument on low inference observations about questioning, it creates a comfortable environment for observees to open up and critically engage with many broader elements of their teaching including participation, planning, and philosophy.

The primary objective of the study is not to study the best practices around questioning, rather it is to use questioning as a point of departure for critical self-reflection by the observed librarians. A secondary objective is to use the observation process as a method to test and refine the instrument’s design. Through the analysis of observee reflections on their teaching and the observation instrument, this study aims to contribute to academic libraries a method by which an observation instrument can promote a peer observation program for professional growth and not scrutiny.

2. Literature review

The study integrates literature from the fields of library and information science (LIS) and education to conduct a comparative analysis of classroom observation approaches. Additionally, this literature guides the study’s approach to question classification, as well as the design of the observation instrument and the pre- and post-observation interviews.

2.1 Observation of instruction

2.1.1 Critical self-reflection

Freire (1998) defines the disposition of good teaching as “right thinking,” which includes teaching by example, being open to multiple answers, and respecting the lived experiences of students. This pedagogy requires that teachers be critical and reflective. However, right thinking is “not a gift from heaven, nor is it to be found in teachers’ guide books, put there by illuminated...
intellectuals who occupy the centre of power. On the contrary, the correct way of thinking that goes beyond the ingenious must be produced by the learners in communion with the teacher responsible for their education” (Freire, 1998, p. 43). Thus the road to good teaching, “right thinking,” runs directly through an iterative, reflective practice, where teachers need to consistently think about the quality of their work, because “thinking critically about practice of today or yesterday, makes possible the improvement of tomorrow’s practice” (Freire, 1998, p. 44).

Bengtsson (2003) rightly points out that while reflection is believed to serve an enlightening function, helping teachers to develop agency, there are two key limitations with the term’s circulation. First, there is ambiguity and uncertainty around “reflection” and thus it is difficult to determine exactly what the process entails. Second, there is a lack of critical examination of the products of “reflection” and what it can realistically achieve.

To address Bengtsson (2003) critique, this study relies on Reale’s (2017) definition of self-reflection for IL instructors, which asserts that it is an intentional process of “questioning all of the assumptions that we have been operating on, and refashioning, reformulating, and reinventing the way we do things” (Reale, 2017, p. 2). Tocco et al. (2023) use the term “pedagogical metacognition” to describe this process. The areas of teaching that require pedagogical metacognition (reflection) are when we detect weakness in our lesson delivery, changes in priorities in higher education, the true impact of our teaching strategies on our students, and the quality of our pedagogical training.

2.1.2 Peer observation
The research demonstrates that peer observation is an effective approach for instructors to engage in self-reflection about their teaching methods. The benefits include developing a supportive network among professionals, fostering critical reflection and dialogue on teaching practices, and improving teachers’ knowledge and skills (Alabi & Weare, 2014; Gordon & McGhee, 2019; O’Leary & Price, 2017).

The literature strongly advocates for a three-part structured approach to peer observation. First, a pre-observation conference where colleagues (observer and observee) establish clear goals and expectations for the observation (Alabi & Weare, 2014; Brewerton, 2004). Second, the subsequent observation, which should focus on observable behaviours rather than the personal characteristics of the instructor to avoid offense or personal criticism (Gordon & McGhee, 2019; O’Leary, 2020). And third, a post-observation conference, which centres around a series of questions to encourage personal reflection on the part of the observee, rather than offering judgments from the observer (Sinkinson, 2011; Whitmore, 2009).

2.1.3 Observation instruments
Where the literature on peer observation programs does conflict is on the question as to whether a peer observation should include an observation instrument. An observation instrument establishes a standardised method for recording classroom phenomena during an observation. Standardised observation instruments are common in the realm of faculty performance, program, and institutional assessment (Hora et al., 2013; Sawada et al., 2002; Smith et al., 2017).

Peer observation programs proposed by Sinkinson (2011) and Snavely and Dewald (2011) eschewed an observation instrument citing their inaccuracy and prescriptiveness. Observation
inaccuracies arise from the Hawthorne effect, wherein individuals under observation alter their behaviour to match perceived expectations (Gordon & McGhee, 2019; McCambridge et al., 2014). To mitigate these issues, O’Leary (2020) suggests using low inference observations in the observation instrument. Low inference observations are factual statements about what was directly seen or heard in the classroom, ostensibly to reduce observer bias. Thus, the present study elected to employ only low inference observations in the design of the observation instrument.

The LIS field has produced few observation protocols. Brewerton (2004) and Middleton (2002) employ checklists of desirable instructional practices. Eastern Kentucky University Libraries (2010) uses a running recording form that asks an observer to list the instructor’s actions and behaviours they found effective and ineffective. However, these three protocols rely on the observer to make determinations about the quality of a lesson rather than fostering reflection on the part of the observee.

Alternatively, Oberlies et al. (2020) adapted three peer observation protocols from outside LIS—Teaching Squares, Teaching Practices Inventory, and Classroom Observation Protocol for Undergraduate STEM—to IL instruction. All three “use observation and reflection to evaluate teaching practices using data and value-neutral feedback” (Oberlies et al., 2020, p. 7).

Teaching Squares involves intervisitation among instructors from four disciplines who focus the post-observation debrief on what they learned about their own teaching from the observation (Haave, 2014). Having modified this protocol to include four librarians, Oberlies et al. (2020) found that participant observers benefited from exposure to new teaching activities they had not previously considered while participant observees pursued critical feedback from observers, which the authors attribute to the protocol’s ability to create “safe, mutually supportive spaces” (Oberlies et al., 2020, p. 24). The Teaching Practices Inventory is a self-administered checklist of practices that invites instructors to reflect on their teaching (Wieman & Gilbert, 2014). Oberlies et al. (2020) redirected the protocol to a single library session and included questions specific to IL instruction. The Classroom Observation Protocol for Undergraduate STEM directs observers to assign standardized activity codes to the instructors (lecturing, conducting a demo, etc.) and students (taking notes, working in groups, etc.) at two-minute intervals during an observation (Smith et al., 2013). Oberlies et al. (2020) streamlined these categories and added IL-specific codes.

2.2 Questioning as pedagogical technique

2.2.1 Questioning
Information literacy requires that instructors create learning environments where students apply critical thinking to how they acquire and use information. Instructor questions are a key factor in achieving this goal. Questioning offers numerous pedagogical benefits, including increased student engagement (Cruickshank et al., 2006; Fusco, 2012; Jacobson & Xu, 2004), enhanced critical thinking (Song, 2019; Walsh & Sattes, 2015), and students’ internalisation and ownership of new information (Grassian & Kaplowitz, 2009). Questioning also serves to conduct formative assessments, which enable instructors to assess student understanding and adapt their teaching methods on the spot (Francis, 2016; Jiang, 2014; Kaplowitz, 2012; Saxton et al., 2018; Walsh & Sattes, 2015, 2011;).
2.2.2 Question classification
Whitver & Lo (2017) conduct one of the few LIS studies on question classification, examining whether instructor questions are scripted or unscripted. This approach does not appear especially fruitful in terms of developing pedagogical competency. Question classification in the education literature tends to focus on the question's underlying purpose, specifically what it requires students to do (Dillon, 1990; Saxton et al., 2018; Walsh & Sattes, 2011).

The education literature has produced an abundance of question classification systems. One question category shared among these systems is the essential question (EQ), which is a thought-provoking and open-ended question that encourages deep thinking (Francis, 2016; McTighe & Wiggins, 2013; Walsh & Sattes, 2011).

From there, various systems exist that differentiate between fact-based questions, which concentrate on recall and developing fundamental knowledge, and thought-provoking questions, which demand students to make judgments or draw inferences (Dillon, 1990; Francis, 2016; Groisser, 1964; Walsh & Sattes, 2015). Saxton et al. (2018) group questions into three main categories: (1) those that elicit information, or draw out prior knowledge; (2) those that shape understanding, or invite students to draw conclusions; and (3) those that press for reflection, or challenge students metacognitively.

2.2.3 Wait time
Questioning typically adheres to the IRF model, consisting of teacher initiation, student response, and teacher feedback (Wood et al., 2017). After each step of IRF, instructors afford silence, called wait time, for students to formulate their responses (Sadker, 2003). Studies into wait time generally focus on the time between initiation and response. Rowe (1974), Paolletti et al. (2018), and Larson and Lovelace (2013) all determine that the ideal wait time to provoke maximum student participation is between three and six seconds. Instructors who feel uneasy with silence and hastily jump in to offer answers in the absence of an immediate response miss a crucial instructional opportunity (Fusco, 2012; Jacobson & Xu, 2004; Kaplowitz, 2012; Walsh & Sattes, 2015, 2011). Ingram and Elliott (2016; 2014) argue that excessive wait time is counterproductive and produces diminishing returns.

2.2.4 Culture and participation
While questions can foment deeper class engagement, instructors should critically reflect on why some students are not participating in class. Dillon (1990) challenges the claimed benefits of questioning, suggesting that they are primarily useful for assessment and behavior control, and may not necessarily enhance students’ cognitive, affective, and expressive processes. Kaplowitz (2012) highlights that verbal questions might not effectively reach all students.

Students bring to class diverse cultural norms and preferences regarding participation (Hicks, 2019). Proficiency in the language of instruction directly influences a student's comfort level in class (Eliason & Turalba, 2019; Tatar, 2005). Yaylcai and Beauvis (2017) found that female students, those with lower English language proficiency, and racial minorities self-reported participating less frequently in class. Instructors should plan for students’ different participation preferences and afford a variety of ways for students to participate that might not include raising a hand to answer a question (Arellano Douglas & Gadsby, 2022; Meyer et al., 2014).
3. Methodology

This study investigates the potential of an observation instrument to facilitate critical self-reflection among library instructors on their teaching practices. To test this, the study employed a research protocol that placed participants in a situation where they were observed while teaching and then asked to reflect on that lesson with the data from the observation instrument in hand. The protocol does not aim to generalise best practices around questioning strategies in library instruction, rather it puts forward a system to foster self-reflection among library instructors.

3.1 Participants

The selection criteria stipulated that participants be members of the library faculty and staff at Queens College, City University of New York, who teach as part of their professional duties and would be teaching during the spring 2023 semester. From the list of library faculty and staff that met the selection criteria (n = 15), I invited five individuals at random. Given my professional relationship with the faculty and staff, I chose to randomly select participants in order to avoid introducing bias into the study about who I believed would be most open to engaging in self-reflection. All five of those invited agreed to participate.

The participants included four women and one man, all who hold a MLS and have between one and twenty-five years’ experience teaching in academic libraries. Four participants reported having formal pedagogical training, which were dedicated pedagogy courses in non-LIS graduate programs. All participants had previous informal training as teachers, principally the experience of shadowing colleagues while they taught. All but one participant had been previously observed while teaching.

3.2 Protocol design

The research protocol adapted the bookended peer observation model proposed by Gordon and McGhee (2019), O’Leary and Price (2017), and Alabi and Weare (2014). Thus, I interacted with participants three times over the course of the study: during the pre-observation interview, observation, and post-observation interview.

First contact was in the pre-observation interview, where I conducted a brief structured interview. Participants were asked to describe their formal and informal instructional training and their experiences being observed while teaching. Two Likert scale questions asked participants about their comfort level being observed for this study and their interest in reviewing data from the observation instrument (see Appendix A). Both Likert questions were repeated in the post-observation interview.

McCambridge et al. (2014) demonstrate that while there are many factors that cause observees to modify their behaviour while under observation, little is actually known about the conditions that give rise to and the magnitude to which the so-called “Hawthorne Effect” influences participant behaviour. Thus, out of an abundance of caution, participants were not shown the observation instrument in the pre-observation interview to mitigate conscious or unconscious adaptation of their teaching style.
The second contact between participants and me occurred during the observation. I sat in each class for 60 minutes and completed the observation instrument on paper. At the end of the observation, I took a few contextual notes on the student-teacher dynamic and topic of instruction to aid my memory. Participants did not see these notes. The types of classes ranged from one-shot sessions, a workshop for a first-year writing course, and semester-long, credit bearing courses. Class sizes ranged from 10 to 16 students and included both undergraduate and graduate students. The variety of teaching contexts is immaterial to the hypothesis of this study. The purpose of collecting classroom observations was not to research teaching methods, which is contingent on student population, but rather to use the observation instrument in an authentic classroom experience and test whether it could elicit meaningful self-reflection on the part of the participant in the post-observation conference.

The third and final contact between participants and me was in the post-observation interview, which was conducted as a semi-structured interview with 19 open ended questions (see Appendix B). The interview began by asking the participants how they felt the class went, how they planned the class, and how they felt being observed. Then I showed participants the observation instrument and provided them with an explanation of how each data point was collected. For each section - types of questions, participation, and wait time - I identified the related data point in the observation instrument and asked the participants a few reflection questions. The final five questions of the post-observation interview asked participants to, in a sense, stand next to me and critically evaluate the design of the observation instrument. I asked participants to comment on whether the data points provided useful information as they reflected on their lesson.

3.3 Observation instrument

The observation instrument (see Appendix C) draws heavily from the research discussed in the literature review section on how to collect low inference observations on instructor questioning methods. Fields 1-6 of the instrument are basic information about location, time, and class of the observation. Fields 7 and 8 work together to record how many unique students participated. In other words, how many students spoke at least once, either answering a question or posing a question.

I recorded every question posed by the participant in the first column of the Question Record table (see Appendix C, field 11). If a question was repeated verbatim shortly after initially being posed, the question was recorded once. If the participant rephrased the question, both forms of the question were recorded as separate questions. Groisser’s (1964) “tugging” or “unproductive questions” such as “what else?” or “does that make any sense?” were recorded.

As I recorded a participants’ questions, I simultaneously counted (one thousand, two one thousand) how many seconds elapsed between the end of the question and a student response and recorded the duration next to each question under the Wait Time column. Time also stopped and was recorded if the participant decided to rephrase, pose a new question, or move on. After the observation, I calculated the average wait time across all questions and recorded the value in field 9. Wait time was selected as a dimension of instructor questioning based on the foundational work of Rowe (1974) and developed by Walsh and Sattes (2015, 2011), Kaplowitz (2012), Fusco (2012), and Jacobson and Xu (2004).
Next in the Question Record table, I recorded yes (Y) or no (N) as to whether a student responded to the question under the Student Response column. Under the Detailed Response column, I would record a no (N) if a student replied with a simple a yes, no, or a single word answer. I would record yes (Y) if the student’s response included some detail beyond the single word response. These dimensions were adapted from observation instruments developed by Paoletti et al. (2018) and Larson and Lovelace (2013), both of which measured the rate at which students responded to different question categories as well as the depth of detail of those responses.

After the observation, I classified each question in the Question Record table (Appendix C, field 11) as Informational (I), if it solicited a specific fact or piece of information; a Check for Understanding (C), if it asked students to describe a concept from the class or voice any confusion; or Analytical (A), if it asked students to evaluate or critique a concept at hand. Each question received only one classification category. These three classes are a composite of the question classification systems discussed in the literature review. However, they are ultimately the original design of this study. Additionally, from a purely ease of use perspective, three classes are a manageable number, and thus useful, for observers during the classification process and for observees during the interpretation process in the post-conference interview.

After classifying each question, I calculated the total number of questions per classification category, the average student response rate for all questions per classification category, and the rate at which student responses were detailed per classification category. These calculations were recorded in field 10 of the instrument.

3.4 Data analysis

I used thematic content analysis to analyse participant reflections on their personal observation instrument. Transcripts of the post-conference interviews were divided into thematic clusters around each data point of the observation instrument as well as compared to established instructional practices in the education and LIS literature.

3.5 Ethical considerations

Classroom observations play a part in determining reappointment, tenure, and promotion in higher education and have real bearing on the future employment of library faculty and staff. Moreover, discussing past classroom observations may cause discomfort if the participant had a previously negative performance assessment. As part of the informed consent process, I shared with participants that all notes and observation instruments would never be disclosed to institutional administrators. The Queens College Institutional Review Board deemed that the research design (2022-0676-QC) posed minimal risk to participants and granted exemption from ongoing oversight.

3.6 Limitations

Two principal limitations arise from this research design. First is that of limited generalisability. The sample size and research setting at Queens College mean that the findings and conclusions drawn from this study may not be applicable to a broader population or different educational settings. However, the purpose of this study is to determine the viability of this particular observation instrument design. The second limitation is that the observation
instrument focused on instructor questioning methods at the expense of collecting data on the full complexity of teaching and learning dynamics in the classroom. However, designing an instrument that even attempts to capture these nuances would be so large and unwieldy that it would create an overwhelming amount of data. The purpose of the instrument is to be a point of departure for self-reflection and a more wide-ranging set of observations would weigh down the process to such an extent as to render it useless.

4. Results

In the following section, I present the combined results from the observation instruments used in all five observations and the participants’ reflections from the post-observation interviews. In these interviews, I first asked participants to reflect on what the low-inference observations reveal about their instruction. I then asked participants to provide feedback on the overall design and usefulness of each section of the observation instrument. In direct quotes from participants, I exclude filler words in order to succinctly convey participants’ opinions. I also attribute the “they” pronoun to all participants to protect their identities.

4.1 Comfort with observation

For participants to engage in meaningful reflection, it is crucial that the observation protocol creates a comfortable environment that encourages self-reflection. During pre- and post-observation interviews, I asked participants the following Likert scale question: On a scale of 1-5, with 1 being very uncomfortable and 5 being very comfortable, how do you feel about someone observing your class? When I repeated this question in the post-observation interview, participants did not see their previous responses. As table 1 shows, their responses to the question both before and after observation were nearly identical, save for participant 4 (P4) who expressed more comfortability post-observation.

Table 1: On a scale of 1-5, with 1 being very uncomfortable and 5 being very comfortable, how do you feel about someone observing your class?

<table>
<thead>
<tr>
<th>Participant</th>
<th>Pre-observation interview rating</th>
<th>Post-observation interview rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>P2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>P3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>P4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>P5</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

4.2 Question classification

Questioning was a common instructional strategy among participants. Table 2 presents the number of informational, checks for understanding, and analytical questions posed by each of the five participants during their observation. On average, participants posed 14 questions per 60-min observation period. There was substantial variance in the number of questions posed by individual participants. Three participants asked between 12 and 16 questions. However, one participant (P2) asked only 2 questions while another (P1) asked 26.
Table 2: Number of questions per question class per participant

<table>
<thead>
<tr>
<th>Participant</th>
<th>Question Classes</th>
<th>Total questions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Informational</td>
<td>Checks for understanding</td>
</tr>
<tr>
<td>P1</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>P2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>P3</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>P4</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>P5</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>Average</td>
<td>7.3</td>
<td>2</td>
</tr>
</tbody>
</table>

Question examples from the observations included:

Informational questions:
“Where have you turned to conduct research in the past?”

Checks for understanding questions:
“What does the author mean in this section?”, “What are two options available to filter our results here?” and “What is the difference between a database and a search engine?”

And analytical questions:
“What do we notice about our results?”

During post-observation interviews, participants reviewed the count for each question type on their observation instrument. I asked participants if they felt that the count for each question accurately reflected their objectives for the lesson. P2 answered in the negative because they were only able to pose two questions (see Table 2) during the observation. Students in fact peppered P2 with questions for the entirety of the 60min observation. P1 and P5 reported that the question counts did align with their objectives. P4 said that “I wouldn’t ask more analytical questions, but maybe more checks for understanding to see if students were following.”

When asked how they decide when to pose a question to the class, P1 and P4 said they do not plan questions while P1 said that they ask questions when they get to the end of a slide. P2, P3, and P5 set aside distinct moments to pose questions. Participants gave examples such as a question to start off the class or questions that precede or conclude an activity. Only two participants reported writing down their questions before class.

When asked to evaluate the question classification system, specifically whether the three categories reflect the types of questions librarians ask, four of the participants agreed that they do. One said they are not prepared to say and would require more information. P3 and P5 said that the three categories generally reflect how they already assess student learning when they use questions as formative assessments during instruction.
4.3 Student response rate and response detail rate

Table 3 presents the count, response rate, and detail rate for each question category of each participant. Count reflects the number of questions asked by the participant in each question category, building on the information provided in Table 2. The response rate represents the percentage of questions posed by the instructor that were answered by students. Finally, the detail rate indicates the proportion of those answered questions where students provided an answer beyond a simple yes, no, or other one-word response.

Table 3: Count, response rate, and detail rate per question category per participant

<table>
<thead>
<tr>
<th>Participant</th>
<th>Informational questions</th>
<th>Checks for understanding questions</th>
<th>Analytical questions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Count</td>
<td>Response rate (%)</td>
<td>Detail rate (%)</td>
</tr>
<tr>
<td>P1</td>
<td>14</td>
<td>86</td>
<td>58</td>
</tr>
<tr>
<td>P2</td>
<td>1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>P3</td>
<td>11</td>
<td>100</td>
<td>55</td>
</tr>
<tr>
<td>P4</td>
<td>9</td>
<td>66</td>
<td>83</td>
</tr>
<tr>
<td>P5</td>
<td>8</td>
<td>100</td>
<td>38</td>
</tr>
<tr>
<td>Average</td>
<td>7.3</td>
<td>91.5</td>
<td>69</td>
</tr>
</tbody>
</table>

P1, who asked the highest number of informational questions, noted that while the response rate was high, the detail rate was less than they would have liked. P1 reflected that “I feel giving a lot of information and then saying, ‘any questions?’ didn’t yield a lot.” P5, who posed three analytical questions and received a response to only one, stated that “I think analytical is harder: you have to drag it out of them. They are responding more to checks for understanding.”

Meanwhile, P4, who asked one analytical question which received a detailed response from a student, said that “I feel with fewer analytical questions, I got better response rate. I am pleased it was not yes/no. I’m asking questions that elicit responses.” P3 expressed dissatisfaction with the response rate and detail rate to their questions.

When asked to critique the collection of this data (do you feel it is useful to see the response and detail response rates for your questions?) participant responses were mixed. P1, P2, and P3 found the detail rate helpful. P1 commented that “it shows depth of material and success for me in what I’m trying to do.” P2 echoed this sentiment: “I want to ask open ended questions. Detailed responses show as achieving that goal and evidence of better learning outcomes.”

P5 expressed a different perspective:

Response rate is more helpful than detail. If I’m asking a lot I’d like to know if they are responding. I need more info on detail rate, but I don’t know how that will help me in planning. Detail expects more rich info, but I’m not sure of the quality of detailed responses from looking at these figures.
4.4 Wait time

Table 4 calculates the average wait time across all questions, answered or unanswered, posed by participants during the observation. I showed participants their average wait time as well as the wait time per question during the post-observation interview.

Table 4: Average wait time across all questions per participant

<table>
<thead>
<tr>
<th>Participant</th>
<th>Average wait time</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>1.9</td>
</tr>
<tr>
<td>P2</td>
<td>2.5</td>
</tr>
<tr>
<td>P3</td>
<td>2.4</td>
</tr>
<tr>
<td>P4</td>
<td>2.0</td>
</tr>
<tr>
<td>P5</td>
<td>4.3</td>
</tr>
<tr>
<td>Average</td>
<td>2.8</td>
</tr>
</tbody>
</table>

Three participants reported that they consciously give wait time to students after posing a question. P5, who gave the longest wait time on average, explained that:

I give them time to respond. Without it, the chances of a response are low. Let them accept a question is out there and feel that they should respond. It should be around three seconds.

P2 acknowledged that “it’s very important,” but does not explicitly give wait time because “anxiety and excitement color my sense of time.” P1 shared this sentiment:

I feel like I should give more [time]. I have a tendency to jump in to help. Time is moving faster and faster in my head. I don’t revel in silence. Silence is a question itself.

Four participants reported that it was useful to see how much wait time they give to students. The fifth participant qualified their response saying that while measure of wait time is useful, it is not as helpful as other data collected during the observation. To them, it was more useful to see wait time per question in the Question Record as opposed to average wait time (see Appendix C, section 11).

4.5 Student participation

Table 5 presents the student participation rate across the five observations. Student participation rate is defined as the number of unique student participants during the 60min observation period. Any student who participated at least once during this period is considered a unique participant.
Table 5: Student participation rate across five observations

<table>
<thead>
<tr>
<th>Participant</th>
<th>Number of students present</th>
<th>Number of unique student participants</th>
<th>Student participation rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>12</td>
<td>10</td>
<td>83.3</td>
</tr>
<tr>
<td>P2</td>
<td>10</td>
<td>7</td>
<td>70.0</td>
</tr>
<tr>
<td>P3</td>
<td>16</td>
<td>5</td>
<td>31.3</td>
</tr>
<tr>
<td>P4</td>
<td>11</td>
<td>3</td>
<td>27.3</td>
</tr>
<tr>
<td>P5</td>
<td>13</td>
<td>8</td>
<td>61.5</td>
</tr>
<tr>
<td>Average</td>
<td>12.5</td>
<td>5.8</td>
<td>47.5</td>
</tr>
</tbody>
</table>

When asked if they were surprised by the student participation rate during their observation, all participants replied that they were not. P3’s comment summarises the responses from the other participants: “I was aware of who was engaging.” However, when asked if they believe students were comfortable participating in the lesson, participant reflections varied. Only P2 and P5 answered that they believed that students were comfortable participating in the lesson. P3 and P4, with the lowest student participation rates, reflected that they had let some students “monopolise conversations.” When a small group dominates the conversation, P4 reports that: “Others feel there is no need to participate or that there’s no room to break in.”

When asked if they thought it was useful to see the number of unique participants, P3 said: “It’s helpful to be aware of the difference between a lot of words and a lot of people speaking words.” P5 repeated this idea and developed it further:

> Its good to know if all students are participating or not. It’s hard to note during the class. I want to make a conscious effort to bring people in. I should look out for others to speak.

P1 acknowledged the potential of this data point to enable self-reflection: “If it were lower, it could be a warning, could show a problem in the clarity of my questions.”

5. Discussion

The use of low inference observations in conjunction with open-ended questions prompted participants to engage in critical self-reflection of their instructional practices. This key finding is largely consistent with the proposed benefits of peer-observation programs in the literature (Alabi & Weare 2014; Gordon & McGhee, 2019; O’Leary, 2020; O’Leary & Price 2017; Sinkinson 2011; and Whitemore 2009). Each participant considered the data presented to them on the observation instrument and related it to their memory about the class or their general teaching philosophy.

5.1 Comfort engaging in observations

Overall, participants were somewhat to very comfortable with the observation experience, as they self-reported to the Likert-scale question in both the pre- and post-observation interviews. The participants’ comfort level can significantly impact their feeling of psychological safety and openness to engage in a conversation about ways they can improve their teaching methods.
was a positive sign that their self-reported comfort levels did not drop as a result of the observation experience (see Table 1).

Care should be taken not to attribute too much weight to any one specific reason why participants were willing to engage critically with data from the observation instrument. The truth lies somewhere among the confluence of factors: (a) They are dedicated professionals who take their work with students seriously; (b) by offering low inference observations, and thus provided feedback on observable behaviour rather than the personal characteristics of the instructor (Gordon & McGhee, 2019), participants were not offended or intimidated by the instrument; or (c) perhaps my previous, positive working relationship with the participants created a trusting environment. I had only worked with participants for six months prior to the start of the study.

5.2 Reflections on student participation

The observation instrument placed significant emphasis on student participation. Participants acknowledged that it is challenging to obtain an accurate assessment of participation levels while teaching. The inclusion of the student participation rate on the instrument aimed to encourage participants to reflect on student engagement across various dimensions, such as race, ethnicity, language, and gender. However, when participants were asked about student comfort level in participating during the observed lesson, their responses delved into other dynamics, such as their perceived approachability or their ability to manage conversations constructively, especially when one student monopolises the discussion. Only one participant (P5) approached the relationship between student background and participation rate, expressing their intention to involve a wide range of students in the classroom.

It was tempting to provide statistics of student participation rate based on their perceived race, ethnicity, or gender. Whitver and Lo (2017) and Shah et al. (2016) investigate participation disparities along racial, ethnic, and gender dimensions by assuming student identities in their methodologies. Ultimately, this study rejected such a methodology because it is impractical, given the diversity of the Queens College student body, and problematic, as it would necessitate making assumptions and generalisations about significant differences within populations (Espinosa de los Monteros & Longmeier, 2022). The literature points out that students who are female, have a lower English language proficiency, or belong to a racial minority self-report lower participation rate in class (Park, 2018; Yaylcai & Beauvis, 2017). The observation instrument aimed to prompt deeper reflections and discussions on inclusive teaching practices, but the race, ethnicity, language, or gender of students did not appear in post-observation interviews.

5.3 Reflections on questioning methods

All participants commented that questioning is a major instructional strategy essential to their teaching philosophy. This is an important distinction when compared to Whitver and Lo (2017), who alarmingly found that most of the library instruction sessions they observed offered zero questions to students. Participants in the present study echoed the benefits of questioning established in the literature: increased student engagement (Cruickshank et al., 2006; Jacobson & Xu, 2004), more thoughtful classroom discussions (Fusco, 2012; Walsh & Sattes, 2015), and the internalisation and ownership of new information (Grassian & Kaplowitz, 2009; Song, 2019). All participants concluded that they needed to increase the number of checks for understanding questions in order to better gauge content acquisition and student engagement. The literature
supports this conclusion as these are the key advantages of incorporating formative assessments in lesson planning (Francis, 2016; Jiang, 2014; Kaplowitz, 2012; Saxton et al., 2018; Walsh & Sattes, 2015, 2011).

More questions are not necessarily better, and the participants reflected that the timing of questions produced better results. As reported in the results section, P4, who asked one analytical question which received a detailed response, said that:

I feel with fewer analytical questions, I got better response rate. I am pleased it was not yes/no. I’m asking questions that elicit responses.

Larson and Lovelace (2013) found that instructors “who carefully crafted a few pointed questions at critical junctures throughout the lecture appeared to be more effective [at obtaining student responses] than those who bombarded the students with questions to stimulate superficial engagement” (p. 116).

Participants’ reflections also provided meaningful insights that were not previously explored in the literature. All participants, except for P5, expressed a desire to incorporate more analytical questions into their instructional approach. The participants’ desire to incorporate more analytical questions likely stems from their recognition of these questions as essential to the complex and intriguing topics covered in library instruction. Instructors are interested in unpacking these concepts with students to promote deeper understanding and critical thinking. The motivation to increase analytical question counts was particularly relevant for P2 and P3 (as shown in Table 2) since they did not utilise analytical questions during the observation. On the other hand, for P1 and P5, who did include 4 and 3 analytical questions respectively, increasing the number further might ignore the type of concept development that builds a class up to effectively engage with such questions.

The importance of properly timing analytical questions within the instructional process was evident in participants’ reflections. As supported by Francis (2016), McTighe & Wiggins (2013), and Walsh & Sattes (2011), analytical questions, akin to essential questions, are best employed toward the end of a lesson. This strategic use of analytical questions aids in promoting deeper understanding and critical thinking among students.

P2’s observation was unusual because they posed only two questions. This might imply P2 presented a lecture-heavy session on account of the low number of questions asked. However, the reality was that students were actively engaged, coming prepared with their own questions for the librarian. From the perspective of this author, this scenario exemplifies an ideal teaching situation, where students actively steer the learning process through their inquisitiveness and interests.

Wait time is an essential part of reflecting on questioning (Fusco, 2012; Jacobson & Xu, 2004; Kaplowitz, 2012; Rowe, 1974; Walsh & Sattes, 2015, 2011). Larson and Lovelace (2013) argue that four to six seconds is the optimal wait time. Participants were outside the mark, averaging 2.8 seconds of wait time and only one participant made it within the optimal wait time. One participant remarked that they preferred seeing wait time measurements next to each question rather than their average time. This would be especially helpful to understand, as an instructor, how your wait time shifts by question type or section of the lesson. This is reinforced by the
case of P1, who had the shortest wait time average but the highest response and detail rates. The fact is that students answered their questions fairly quickly.

While P1 also had a high student participation rate, instructors should be careful when students answer questions too quickly as this can lead to a small group of students dominating the conversation before others have the chance to participate. This was the case with P4, who had a very short wait time and a low student participation rate.

5.4 Critique of the observation instrument and protocol

The secondary objective of the post-observation interviews was to elicit feedback from participants regarding the instrument’s design. When asked about the usefulness of examining questions as a point of reflection for library instructors, participants unanimously acknowledged the value of questions as essential tools for effective instruction. They emphasised that questions play a vital role in actively engaging students, moving away from a one-sided lecture format, and fostering critical thinking. For example, P1 expressed concern about information-heavy classes and highlighted how questions empower the class to transcend the traditional "come, listen to me, leave" approach. Similarly, P4 recognized questions as pivotal moments to involve students in the learning process.

Participants also expressed appreciation for a dedicated space that focused on their individual choices as instructors. Through this introspection, they were able to identify areas for improvement, including the incorporation of various engagement techniques. P3’s comments aptly encapsulated this sentiment, stating:

It’s been a while since I revised my script. Seeing space for me to incorporate other checks and engagement is important for my practice.

The protocol shielded participants from the instrument until the post-observation interview to minimise the Hawthorne effect. The purpose of this was to put the instrument to the test and see if it could produce critical self-reflection without suggestion from the observer. However, in a true peer observation program, where a library instruction coordinator would elect to use the instrument as part of professional development, the pre-observation interview should encourage open dialogue between the observer and observee (Alabi & Weare, 2014). This should include disclosure of the instrument. Besides, shielding participants from the instrument would only work once before everyone understands its focus on questioning. As the literature has demonstrated, asking good questions is sound pedagogy, and the instrument’s goal is to prompt instructors to think about their use of questions more intentionally and to see growth in the use of questioning techniques.

6. Conclusion and implications for future research

The findings of this study highlight the effectiveness of low-inference observations and open-ended questions in fostering critical self-reflection among participants. The participants felt sufficiently at ease with the observation process, leading to fruitful discussions on how to refine questioning strategies and improve student participation.

Although participants were candid about student engagement in their class, they were reluctant to deeply consider how other factors, such as race, ethnicity, language, or gender, contribute to
student participation. Many participants approached this issue from a perspective of self-blame, attributing it to deficiencies in their own teaching methods, wait time, or question formulation, without addressing potential systemic barriers.

Future research should investigate the next step in a peer observation program. One possibility is to conduct a second round of observations to evaluate participants’ progress in their utilisation of questioning techniques. Additionally, introducing formal professional development on question asking after the initial round of observations could be beneficial. The second observation could measure the growth in basic questioning techniques as well as encourage reflection on the influence of race, ethnicity, language, gender, and teacher positionality in the classroom. Despite the absence of these topics in the current study’s post-observation conferences, the low-inference approach did prove successful in creating an environment where the observee actively engaged in discussions, making the post-observation interview a participant-led exchange of ideas rather than an imposition of the observer’s teaching philosophy.

To further enhance the validity of these findings, future studies should employ the observation instrument with an observer who is not personally or professionally acquainted with the participants. This approach will help evaluate whether the low-inference and open-ended questions genuinely facilitate a comfortable space for participants to engage in critical self-reflection without any potential biases or preconceptions.

In conclusion, this study underscores the value of low-inference observations and open-ended questions in peer-observation programs. The observation instrument does encourage library faculty and staff to reflect critically on their teaching practices and explore student-centred strategies.

Declarations

Ethics approval
The Institutional Review Board at Queens College, City University of New York deemed that this research design (2022-0676-QC) posed minimal risk to participants and granted exemption from ongoing oversight.

Funding
Not applicable.

AI-generated content
No AI tools were used.

References


Francis, E. M. (2016). Now that’s a good question! How to promote cognitive rigor through classroom questioning. ASCD.


**Appendix A**

**Pre-Observation Interview**

1. What formal instructional training do you have (ex: part of a degree program, workshop, professional development)?

2. What informal instruction training do you have (ex: shadowing colleagues, readings, etc)?

3. Has your teaching ever been observed in the past?

4. Can you tell about that experience and purpose of the observation: How you felt about it? What were the ultimate outcomes?
5. On a scale of 1-5 with 1 being very uncomfortable about someone watching your class and 5 being very comfortable about someone observing your class, how do you feel going into the observation?

6. On a scale of 1-6 with 1 being not interested at all in the outcomes of the observation and 5 being very interested in the results of the observation, how do you feel going into the observation?

Appendix B

Post-Observation Interview

Part I: Initial Feelings about the Observation
1. Tell me about how you felt that the class went.
2. Tell me about how you planned this class.
3. What is your definition of student participation and what do you do in lesson planning or delivery to encourage student participation?
4. On a scale of 1-5 with 1 being very uncomfortable about someone watching your class and 5 being very comfortable about someone observing your class, how did you feel about being observed?
5. What could have made you feel more comfortable?

Part II: Observation Instrument Results

A. Types of Questions
6. You asked X informational, Y checks for understanding, and Z analytical questions during your lesson. Do you feel that accurately reflects your objectives for the lessons?
7. How do you decide when to pose a question to the class?
8. Did you plan out any questions for this class? If so, what were they? Why did you plan these questions?

B. Participation
9. X out of Y students present participated in your class. Does this figure surprise you?
10. You have an X% response rate to your questions. Does this figure surprise you?
11. Of questions that students responded to, Y% were detailed responses. Does this figure surprise you?
12. Do you feel that students were comfortable participating in the lesson I observed, why or why not?
13. What could you do in your future lesson planning or delivery to increase student participation?

C. Wait Time
14. On average, you gave students X seconds to respond to your questions. Do you consciously give wait time? How do you feel about giving wait time to students?

Part III: Critique of the Instrument
15. Do you feel it is useful to see the number of unique participants out of total number of students?
16. Do you feel it is useful to see the response and detailed response rates for your questions?
17. Do you feel it is useful to see how much wait time you give to students?
18. What do you think about these different types of question codes? Do you think these three categories reflect the types of questions librarians ask?
19. As a whole, do you feel that looking at questions is a useful point of reflection for library instructors? Why or why not?

Appendix C

Observation Instrument

1. Date: ________________________ 2. Location: ______________________________
3. Instructor: _______________________________ 4. Course: _______________________

10. Question Coding:
   - Informational (I) Total: ______ Resp. Rate: ______ Detail Rate: ______
     Solicits a fact or piece of information from students.
   - Check for Understanding (C) Total: ______ Resp. Rate: ______ Detail Rate: ______
     Asks students to describe a concept from the class in their own words.
   - Analytical (A) Total: ______ Resp. Rate: ______ Detail Rate: ______
     Asks students to evaluate an activity or to draw conclusions about a concept.

11. Question Record*

<table>
<thead>
<tr>
<th>Question</th>
<th>Wait Time (s)</th>
<th>Student Response (Y/N)</th>
<th>Detailed Response (Y/N)</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Full page of rows provided in actual instrument