Holistic and Iterative Development and Dissemination of Conceptual Traffic Signal Questions

Masoud Ghodrat Abadi, S.M.ASCE[1]; David S. Hurwitz, A.M.ASCE[2]; and Shane Brown, M.ASCE[3]

Abstract: Conceptual questions can be used to improve student's understanding through interactive engagement and formative assessment; however, there is a lack of developed conceptual questions in transportation engineering. This paper reports on the development, implementation, refinement, and evaluation of conceptual questions about traffic signals that include students rating of confidence in their solution. Based on student and practicing engineer traffic signal misconceptions, 94 conceptual questions were developed and implemented by 10 public university instructors. Five patterns of student responses were identified in terms of correctness and student confidence: all correct, all confident, all lack confidence, correct-confident combined with incorrect-lacks confidence, and correct-lacks confidence combined with incorrect-confident. Furthermore, the experiences of instructors using conceptual questions with students were considered through semi-structured interviews. Conceptual questions about fundamental aspects of traffic engineering were most frequently selected, conventional multiple choice questions were the most popular type of questions, and enriching exam materials or creating challenging discussions were the primary goals of implementation. While every participant expressed that they will use the material again in the future, potential barriers to adoption remain. **DOI: 10.1061/(ASCE)EI.1943-5541.0000289.** © *2016 American Society of Civil Engineers.*

Author keywords: Conceptual questions; Faculty adoption; Transportation engineering curriculum; Transportation engineering education; Clinical interview; Self-efficacy.

Introduction

Significant attention has been paid to the reform of engineering courses in recent years, with millions of dollars invested in this effort. A significant portion of the investment is focused on improving student conceptual knowledge through reform of the engineering classroom. This process involves both pedagogical practices and curricular materials that can be used in these practices. The development of conceptual knowledge is suggested in contemporary publications from learning experts (NRC 1999a, b). Conceptual knowledge is focused on the underlying concepts of a discipline, with the assumption that knowing these fundamental concepts is critical to applied understandings in the field. Classroom reform efforts are largely focused on increasing the level and quality of active learning. Michelene Chi, a noted researcher in this area, developed the ICAP hypothesis, which suggests Interactive, Constructive, Active, and Passive Learning environments are decreasingly effective (Chi 2009). An example of passive learning is taking notes without any substantial engagement in that process. In contrast, an example of interactive learning is groups of students working on problems together that are harder than they can solve individually.

For students to develop conceptual understanding through interactive engagement, conceptual problems are required. Conceptual questions have been developed in physics (e.g., O'Kuma et al. 2003) and mechanics of materials (Brown and Poor 2010), but are largely nonexistent for traffic signals, or transportation engineering more broadly. A review of three highly regarded traffic and transportation engineering textbooks reveals that much attention is paid to design and calculation problems rather than conceptual questions. For instance, while chapters on either intersection control or signal timing and design in three major textbooks presents more than 100 problems, less than 15% of these problems could be categorized as conceptual questions [Traffic Engineering, Roess et al. (2011) (15 problems/no conceptual), Traffic and Highway Engineering, Garber and Hoel (2015) (28 problems/14 conceptual), and Highway Engineering and Traffic Analysis, Mannering and Washburn (2013) (58 problems/no conceptual)].

The focus of this paper is then a holistic description of the development, implementation, and adoption of a large set of conceptual exercises related to traffic signal engineering and a study of the relation between correctness and content self-efficacy. Included in the developed problems is an evaluation of content self-efficacy, or a student's confidence that their response is correct (Bandura 1997). This follows previous patterns in concept question development (e.g., Brown and Poor 2010) but add the previously unexplored study of the relationship between problem correctness and self-efficacy. This framework provides some clarity on different levels of activity in the classroom that can be used for conceptual knowledge development and classroom reform efforts. This study is original for it is the first of its kind in transportation engineering.

Background

Formative assessment is central to the promotion of an active learning process. Formative assessment could be interpreted as "encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are

[1]Graduate Research Assistant, Oregon State Univ., 101 Kearney Hall, Corvallis, OR 97331. E-mail: ghodratm@oregonstate.edu

[2]Associate Professor, Oregon State Univ., 101 Kearney Hall, Corvallis, OR 97331 (corresponding author). E-mail: david.hurwitz@oregonstate.edu

[3]Associate Professor, Oregon State Univ., 101 Kearney Hall, Corvallis, OR 97331. E-mail: Shane.Brown@oregonstate.edu

Note. This manuscript was submitted on October 1, 2015; approved on February 17, 2016; published online on April 27, 2016. Discussion period open until September 27, 2016; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Professional Issues in Engineering Education and Practice***, © ASCE, ISSN 1052-3928.

engaged" (Black and Wiliam 1998). Research on formative assessment supports the claim that students receiving feedback from others on their understanding has a positive impact on their learning. It has been consistently agreed that formative assessment is a perfect tool for optimizing teaching efforts while it improves student learning (Yan and Cheng 2015). There are numerous studies that support the effectiveness of formative assessment on students' learning. For example, Chung et al. (2006) attempted to address the issue of improving student learning via an online formative assessment in the form of a computer-based system for circuit problem solving and found overwhelmingly positive outcomes for computer-based discussions. In another study, Ludvigsen et al. (2015) studied the effectiveness of formative feedback practice at large lectures in psychology, using clicker questions. They stated that by keeping students engaged and active, formative feedback practice supports students in learning.

As identified in formative assessment, three elements are interactively engaged in the active learning process: (1) conceptual exercises, (2) students' conceptual understanding, and (3) instructors' efforts. In other words, for students to develop conceptual understanding through an active learning process, sophisticated conceptual exercises should be developed and employed by instructors.

Conceptual exercises are the first element in an active learning process. The development of conceptual questions often is done by instructors in isolation from one another and the absence of data on student misconceptions. This is a critical oversight, for conceptual questions need to provide the best way for exposure of new information and they must make the prevailing knowledge free of misconceptions. Therefore, in order to address all the critical conceptions in an appropriate set of questions, a holistic and iterative process is required, which of course, could not be done individually. This process begins with recognition of misconceptions and development of conceptual questions through clinical interviews. Clinical interviews are focused on uncovering an individual's way of thinking about an idea, and are based on the assumption that individuals have unique features of their understanding. The clinical interview method with a semistructured protocol allows the interviewer the required flexibility to ask probing questions to elicit individualized meanings in the interview on the basis of interviewees' responses (Hurwitz et al. 2014).

Students' conceptual understanding is the second element in the active learning process. Developing student confidence that they can solve conceptual problems is critical in conjunction with developing conceptual understanding. Self-efficacy is one's belief that they can successfully execute a task, and was developed by Bandura (1997). A common measure of self-efficacy is asking students to rate their confidence they can solve a particular problem and has been used in ranking tasks in the past (Brown and Poor 2010). Conceptual questions in which students are often wrong but confident they are correct can be the most problematic to overcome. There is an absence of research relating conceptual question response with self-efficacy.

Instructors' effort is the third element in the active learning process. Although designing conceptual exercises has considerable opportunity to enhance learning, broad impacts require utilization by more than just the developers. There is currently an out of balance proportion of development to utilization of novel curriculum. Researchers suggest that the adoption of innovations requires that the innovation has characteristics that make it adoptable, including relative advantage, complexity, and trialibity. Relative advantage is how much better or worse the innovation is compared to the current alternative(s). Complexity refers to how complex the innovation is and trialability refers to how easy it is to try the innovation (Rogers 1995).

Purpose and Research Questions

As previously mentioned, there are very few conceptual exercises related to transportation engineering readily available to faculty. However, such exercises are critical to individual efforts to change classroom practices. This study discusses the development of a large set of conceptual exercises in the traffic signal system domain and holistically addresses student results, and instructors' feedback on use of the developed material. Therefore, the research questions aim to investigate the designed curriculum from two perspectives, those of students responding to and teachers implementing the conceptual questions.

The research goal is to document student responses to the developed conceptual questions and evaluate their confidence in solving conceptual problems. The development and research effort described herein addresses these issues by including a self-efficacy scale on each problem and sharing both the student response results and how they relate to their self-efficacy for that problem. To address this research goal, the following research question was investigated:

- What relationships exist between the correctness and self-efficacy of student response to conceptual questions about traffic signal systems?

Another goal is to evaluate instructors' experience in utilizing the developed material in classrooms and the likelihood of future adoption by them. Although adoptability characteristics were considered in the development process, it is important to understand the implementers' perspectives to allow for informed modifications to existing exercises and development of new exercises. As such, instructors who utilized the exercises were interviewed, in an attempt to understand how they used the exercises and the characteristics of the questions that either encouraged or discouraged future use. Thus, the following research questions were investigated:

- How were the conceptual questions administered by instructors?
- How will the conceptual questions be used by instructors in the future?
- What were the perceived strengths and weaknesses of the conceptual questions?

Methodology

This effort includes the development, dissemination, use, and evaluation of conceptual questions related to traffic signals. The evaluation includes both student and instructor data, in alignment with the aforementioned research questions. The development and dissemination of the conceptual questions is described in the Results section. Methodologies associated with student and instructor data are described in the following.

Conceptual Question Development

Based on an examination of student and practicing engineer traffic signal system misconceptions (Hurwitz et al. 2014), 94 conceptual questions were developed, all of which were reviewed by external academic and professional transportation engineers who are experts in traffic signal systems and revised by the project team, which included the authors of this paper. The six question topics included yellow change intervals, red clearance intervals, cycle length, coordinated signals, effective green time, and actuated signals. The developed curriculum for each topic included three types of conceptual questions:

- Interpretation questions in which a quote is presented, and students must describe in their own words if the quote is correct or not, and why (Table 1);

Table 1. Example Interpretation Questions for Each Topic

Concept	Interpretation questions
Actuated signal	"Actuated traffic signals work better when processing low-volume traffic conditions, because in high-volume traffic conditions the signal operates more like a pre-timed traffic signal."
Coordinated signal	"Coordinated signals along a corridor or in a network are always pre-timed, they cannot be actuated."
Red clearance interval	"Intersections with higher traffic volumes should have a longer red clearance interval time to ensure that vehicles entering the intersection at the end of the yellow change interval can clear the intersection."
Yellow change interval	"The yellow change interval is used to clear the intersection of vehicles and pedestrians."
Cycle length	"Volume is the only factor that affects the cycle length for a pre-timed isolated intersection."
Effective green time	"In an actuated signal, effective green time is the actual duration for which the signal is green for a particular movement."

- Ranking tasks, in which 4 to 5 pictures are presented and students are asked to order the pictures based on a particular variable of interest from highest to lowest (Fig. 1); and
- Multiple choice conceptual questions, questions with one (or more) correct answer and some incorrect distractors. Distractors are misconceptions that were determined from previous study (Fig. 2).

Question Dissemination and Implementation

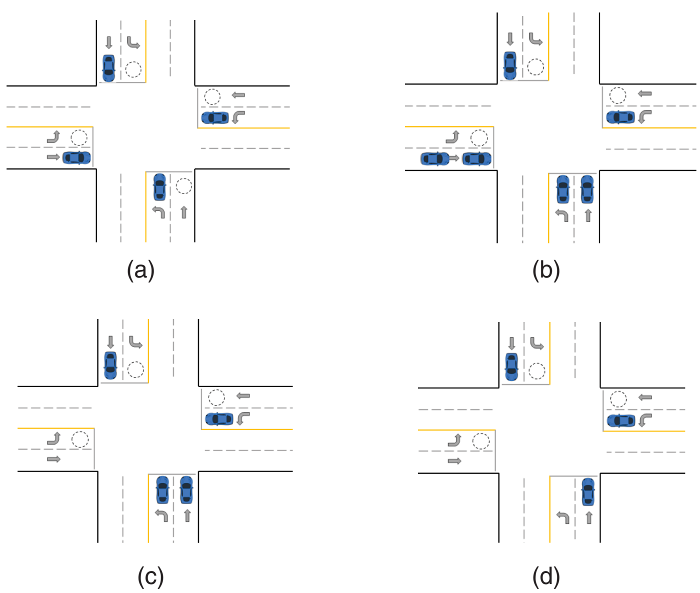
A group of 30 university instructors who were known to the researchers to be engaged in transportation engineering education or interested in methodological improvement in teaching techniques were contacted by email for participation. These potential participants were provided details about the project, including the

Traffic Signal Concept Inventory (TSCI)
Actuated Signal

Ranking Task- 2

The cycle length of a fully-actuated traffic signal can vary based on the traffic demand on the protected phases. The following figures show four different vehicular demand situations for the protected phases at a fully-actuated intersection.

Rank the figures based on the traffic that would create the longest possible cycle length. All concurrent movements can be served by a single signal phase. Assume that the numbers of cars in the figures are a proportional representation of the observed vehicular volumes.



Longest _______ _______ _______ _______ Shortest

Or, the cycle length should be same for all intersection four cases _________

Or, the information is not adequate to compare the cycles _________

How sure are you of your answer? (circle one)

Basically Guessed Sure Very Sure

1 2 3 4 5 6 7 8 9 10

Fig. 1. Example of a ranking task

Traffic Signal Concept Inventory (TSCI)
Actuated Signal

Concept Inventory Question-4

Which of the following statements is correct about the vehicle detection system at actuated signal?

A. Longer detection zones can unnecessarily extend a phase causing delays for other phases
B. Fully actuated signals always require advanced detectors
C. Advanced detectors are installed to measure the speed of approaching vehicles
D. Detectors compute delay at the intersection
E. Length of the detection zone has no effect on phase termination

How sure are you of your answer? (circle one)

Basically Guessed					Sure			Very Sure	
1	2	3	4	5	6	7	8	9	10

Fig. 2. Example of a multiple choice conceptual question

content areas considered, types of questions, methods of implementation, and the required investment of time. Participants were all offered a $200 honorarium for their contribution to the project.

Ultimately, 10 university instructors agreed to participate (33% response rate). All participants were public university instructors from across the country (Fig. 3). Participants included one professor, three assistant professors, and six associate professors. Teaching experience varied among participants with an average of 10.4 years (range from 5 to 29 years).

In total, six question topics were provided to 10 universities. Each topic included three question types, resulting in 730 individual student responses (Table 2). The experiences of instructors using the conceptual questions with students were considered. To do so, researchers scheduled semistructured interviews with each participant. An interview protocol was created over several iterations. An original protocol was developed with the research goals in mind focused mainly on teachers' experience in utilizing and demonstrating the material in their classrooms. A pilot in-person interview was conducted with one of the participants, and the

Fig. 3. Geographical distribution of participants (map data ©2016 Google, INEGI)

Table 2. Concept Areas and Question Types

Concept areas	Multiple choice	Ranking tasks	Interpretation	Number of responses
Yellow change interval	5	5	9	131
Red clearance interval	5	5	8	270
Coordinated signals	4	5	4	0
Actuated signals	5	5	5	20
Effective green time	5	5	5	193
Cycle length	5	4	5	116
Total	29	29	36	730

results were analyzed. The interview protocol, including probing questions, was further developed based on the outcomes of the pilot analysis.

Participants were all contacted via email to schedule the interviews. They were provided with three modes for the interview: Skype, over the phone, or in person. After scheduling the interviews, the main questions, not including the probing questions, from the interview protocol were emailed to participants, making sure that they are aware of the purpose of interview. Eight participants elected interviews over the phone, while two elected an in-person interview. All interviews were audio recorded.

At the initial phase of this project, instructors were asked to pick their desired concept areas and try the developed materials with their students in one of three ways: as an in-class quiz or exam, as a homework assignment, or as in-class formative feedback. After collecting the student responses from each of the implementations, instructors were asked to scan/email the results or hard-copy mail the results to the research team. This way, student data was collected along with possible edits or revisions to the questions and solutions.

Student Data Reduction

Student data were received from instructors in three ways: transcribed in Excel, scanned digitally, or in a hard copy. The instructors generally included information about which questions were used, how the questions were administered, and whether they had made any adjustments to the questions (adding new responses, changing correct responses, etc.). The student responses were transcribed into Excel by one undergraduate and one graduate student independently, and compared. Answers were recorded as correct, partially correct, or wrong along with the reported self-efficacy. A standard data-entry document was created in Excel for each topic area to ensure the consistency of the response entries across the various schools. In addition, nomenclature was established for assessing correctness of responses. Interrater reliability was maintained by identifying conflicts between the two students, which were resolved by in-person meetings where a consensus was reached. Next, scripts were developed in *R version 3.2.0* to analyze the data and generate the visualizations, which were reviewed by the research team to identify cases of interest.

Instructor Data Reduction

Each interview was audio recorded and transcribed into a text file. The transcribed interviews were coded and analyzed using an online qualitative analysis application, *Dedoose version 6.0.24*.

The transcriptions were coded to organize responses based on the primary and probing interview questions. Fifteen major codes and 15 minor codes were developed for this purpose (Table 3). The analysis codes include participants' background and their reason for participation, their approach in implementing the developed

Table 3. Analysis Codes

Major code	Minor code
Years of teaching experience	—
Reason for participation	Justification of participation
Covered topics	Why these topics?
How they used the material?	—
In what class?	Course type
	Course name
	Course level
	Number of students
Best type of class for material	Why this class is best?
Is it an effective learning tool?	—
Preferred type of question	Why?
Specific question you like	Why?
Changed way of teaching	Why?
Future use	How?
	Modifications
	Use it again
Introduce it to colleague?	Summarize the project
Interesting facts about the materials	—
Deficiency	Detail of problem
Interesting quotes	—

material in classrooms, and participants' opinion about the designed curriculum. Additionally, codes highlighted if participants were eager to use the materials in future, if they are willing to disseminate it by introducing to colleagues, and if they identified any deficiencies that could hinder further dissemination of material. A separate code was used to capture any quotes from interviewees, which could be helpful in examining the designed curriculum. Different codes were combined in order to find specific patterns among the experiences of all participants.

Results

Student Responses

Student responses to the conceptual questions were considered on two independent but related measures, correctness and self-efficacy. In many instances, individual question responses were either correct or incorrect. However, in some instances, such as for interpretation questions, responses were also coded as partially correct. Self-efficacy was self-reported on a 10 anchor rating scale ranging from 1 (not at all confident) to 10 (very confident). Five distinct and meaningful classifications were identified and are described in the following subsections.

Type 1—All Responses Are Correct

The Type 1 classification describes each student providing a correct answer to the conceptual question. This result might indicate one or more of three realities, the question itself is too easy to provide a useful measure of a difficult concept, the material has been conveyed through instruction in a highly effective way, or the population of students responding to the question are of high aptitude. Fig. 4 shows an example of a question with a Type 1 response in which all of the answers are correct. For answers to this question, measures of self-efficacy were distributed across the 10 anchor rating scale, although the largest majority are on the positive side [Fig. 4(a)—Self-efficacy: mean = 8.02 and standard deviation = 2.17]. This question was administered as part of an in-class quiz among 42 students, by an instructor with 12 years of teaching experience.

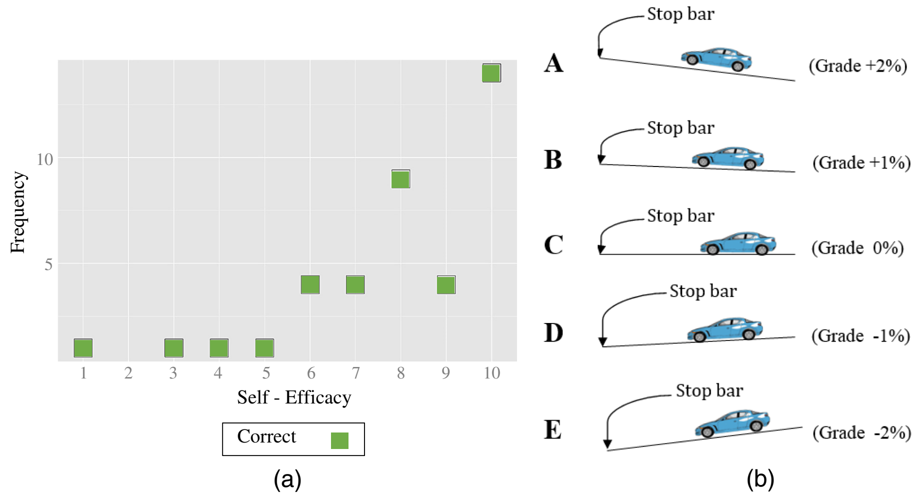


Fig. 4. Example type 1 response

Fig. 4(a) illustrates student responses to a ranking task that dealt with the yellow change interval. In this question, students were asked to rank the duration of yellow change interval for five different intersection approach grades [Fig. 4(b)]. In this case, it seems that everyone is aware of the reverse effect of approach grade on the yellow change interval, although several students are unsure of their correct response.

Type 2—All Responses Are Confident

The Type 2 classification describes situations in which no matter right or wrong, each student is confident about the response, determined to be all confidence ratings of 6/10 or above. This might be representative of a very familiar topic in which all the students have previous knowledge about the concept. Nevertheless, it can be problematic that incorrect answers were aligned with high levels of self-efficacy. Fig. 5 shows an example of a Type 2 response where all students are confident about their answers [Fig. 5(a)—Self-efficacy: mean = 7.52 and standard deviation = 1.29]. This question was administered as part of an in-class quiz among 25 students, by an instructor with 8 years of teaching experience.

Fig. 5(a) illustrates student responses to a multiple choice conceptual question that dealt with the yellow change interval. In this question, students are asked about the purpose of using the yellow change interval [Fig. 5(b)]. In this case, it seems that the question is so straightforward for everyone that they are sure in their answer. However, some are not aware that all the listed items are actually forming the notion of using the yellow change interval.

Type 3—All Responses Lack Confidence

The Type 3 classification describes situations in which no matter right or wrong, each student lacks confidence in their response, defined as all respondents noting a self-efficacy of 6 or below. This might be demonstrating either a very unfamiliar topic or a superficial knowledge about the questioned concept. Instructors who receive these types of responses might consider either expanding their curriculum or presenting more detailed instructions in their classroom regarding the specific topic. Fig. 6 shows an example of a Type 3 response where all the students are doubtful about their answers [Fig. 6(a)—Self-efficacy: mean = 3.80 and standard deviation = 1.38]. This question was administered as a homework assignment to 25 students, by an instructor with 8 years of teaching experience.

Fig. 6(a) depicts student responses to a ranking task that dealt with cycle length. In this question, students are asked about the influence of different control types on the cycle length [Fig. 6(b)]. In this case, it seems that the question is combining different concepts in such a way that students cannot surely rely on their knowledge to select the correct answer. The answer to this question is that enough information is not provided, but it looks like students cannot certainly link all the missing parts to make a correct judgment about cycle length.

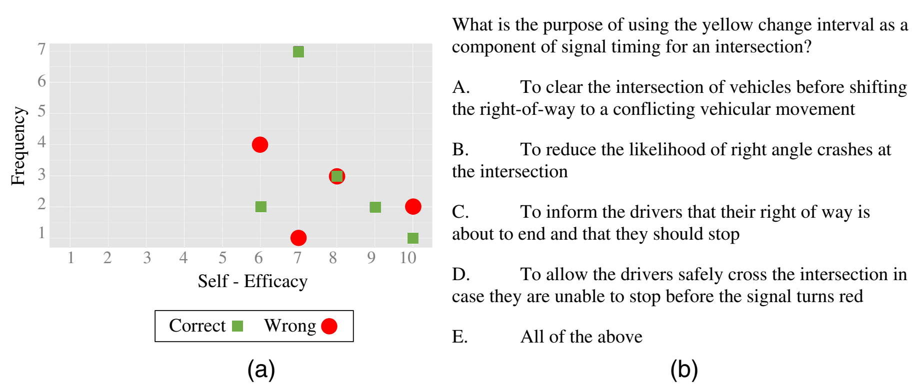


Fig. 5. Example type 2 response

(a)

The following list includes signalized intersections with different types of control and total hourly traffic volumes. Rank them based on the longest cycle length required at each signal.

	Control type	Volume
A	Isolated, Fixed time	1200 veh/hr
B	Isolated, Actuated	1200 veh/hr
C	Coordinated, Fixed time	1200 veh/hr
D	Coordinated, Actuated	1200 veh/hr
E	Isolated, Semi-actuated	1200 veh/hr

(b)

Fig. 6. Example type 3 response

Type 4—Right Is Confident, Wrong Is Doubtful

The Type 4 classification describes one of the most rational cases observed in students' responses, where individuals with right answers have higher self-efficacy values while wrong answers are predominantly made by students who doubt their response. This could be indicating a logical classification in each classroom and about every questioned topic where robust answers are aroused from a solid knowledge and feeble responses are derived from infirm information. Fig. 7 shows a Type 4 response in which right is confident and wrong is doubtful [Fig. 7(a)—Self-efficacy for correct answers: mean = 7.44 and standard deviation = 1.59/Self-efficacy for wrong answers: mean = 4.80 and standard deviation = 1.30]. This question was administered as an in-class quiz to 14 students, by an instructor with 6 years of teaching experience.

Fig. 7(a) shows student responses to a ranking task that dealt with red clearance. In this question, students are asked about the influence of different traffic volumes on the duration of the red clearance interval [Fig. 7(b)]. The key point in this question is that the duration of the red clearance interval is not related to the traffic volume but it is based on the speed and length of vehicles and the geometrical aspects of the intersection. It looks like students who are informed of this fact are sure in choosing the right answers and of course they are right. On the other hand, apparently there are students who do not seem to be aware of the mentioned fact so they are making unsure guesses and of course they are wrong.

Type 5—Wrong Is Confident, Right Is Doubtful

The Type 5 classification describes the most threatening type of responses in which individuals with wrong answers are quite confident that they are right while respondents with correct answers are doubtful of their answers accuracy. Instructors who receive this type of response should certainly dedicate more effort to enrich their instructions in the specific area. Fig. 8 shows a Type 4 response where wrong is confident and right is doubtful (Self-efficacy for correct answers: mean = 4.50 and standard deviation = 0.71/Self-efficacy for partially correct answers: mean = 6.14 and standard deviation = 1.21/Self-efficacy for wrong answers: mean = 7.75 and standard deviation = 1.50). This question was administered as an in-class quiz to 13 students, by an instructor with 7 years of teaching experience.

Fig. 8(a) illustrates student responses to a multiple choice conceptual question that dealt with the actuated signals topic. In this question, students are asked about the cycle length of an isolated actuated intersection [Fig. 8(b)]. In this case, it seems that either students do not have a profound knowledge about the elements of cycle length in actuated signals, or they have received incomplete details about the concept. The right answer here is that there is no fixed cycle length (B), passage time effects the duration of cycle (C), and that the length of the detection zone can affect the duration of cycle (E).

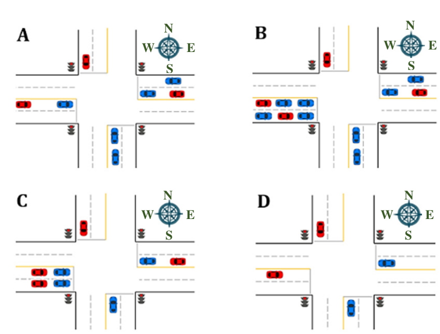
Instructor Perceptions

Instructor perceptions were organized into two categories, their approach in utilizing materials in their classroom and their general impressions of the curriculum's usefulness now and in the future.

Classroom Details

The conceptual questions were implemented in four different types of classes: introduction to transportation engineering (4), advanced

(a)



(b)

Fig. 7. Example type 4 response

(a)

Which of the following statements is correct regarding the cycle length of an isolated actuated intersections?

A. There is a maximum cycle length

B. There is no fixed cycle length, it varies depending on the existing demand

C. Passage time affects the duration of cycle

D. Cycle length can vary based on the phase order at the intersection

E. Length of the detection zone can affect the duration of cycle

F. All of the above

(b)

Fig. 8. Example type 5 response

traffic control (3), urban transportation systems (1), and transportation design (1). Three of these classes were offered at the graduate level while six were offered at the undergraduate level (four in the junior year and two in the senior year). In total, the material was offered in seven required courses and two elective ones.

Participants were consistent in their opinion that a senior/graduate level technical course was the best fit for the material (9/9). Some of the participants also agreed that an introductory undergraduate course was a good choice for the material (4/9).

Preferences for Question Topic and Type

The conceptual question topics were selected by participants with the following frequency: red clearance intervals (7), yellow change intervals (6), cycle length (5), effective green time (4), and actuated signals (1). Participants suggested that the relatively higher frequency of red clearance intervals, yellow clearance intervals, and cycle length was either due to the fact that they were more common components of the transportation engineering classes being offered (4/9) or that they are fundamental concepts in traffic engineering (3/9).

The preferred types of conceptual questions were selected by participants with the following frequency: multiple choice (4), ranking task (3), interpretation (1), no particular preference (1), and context dependent (preferences depend on presented topics) (1). Participant justifications for the preference of multiple choice questions were described in one of the following three ways: (1) they are a common format for an exam question, (2) they are easy to grade, and (3) it is the "right way" to present a problem. Participants who chose the ranking task as their preferred question type supported their reasoning by stating that ranking tasks promoted discussions because it is possible to rank items in various ways, or that they are visually rich so students got insight about the concepts while working with the diagrams and figures.

Preferences for Question Uses

Participants used the conceptual questions in one of four ways with the following frequency: as elements of individual in-class quizzes (8), group discussions (4), homework (3), and exams (2). One participant commented that:

"Those might be the kind of material that you'd start off with at the very beginning of the semester just almost as a review to see where are people starting points, if they already know some of these basic concepts or are there some areas where you need to go back and fill in the few gaps before you moved on."

Indicating that the material could provide formative feedback in the orientation to a content area that the instructor could use to adjust instruction dynamically to the knowledge of the students in a particular class at a particular time.

Implications for Future Teaching Practice

Seven of the nine instructors stated that conceptual questions would not fundamentally change their way of teaching. Justifications for this general statement were most commonly related to not having enough time to make such changes. However, one participant whose teaching method was influenced by the use of this material mentioned that, "Personally, I think about the assessment and evaluation of student understanding in a dramatically different way than I did before I started working on this project."

Additionally, all of the participants indicated that they would use the conceptual questions in their classes in the future, either as is (3/9) or with slight modification to the questions or solutions (7/9). It is difficult to categorize how participants will use the material in the future, for a wide range of possibilities were proposed. Future use most commonly depended on how the material was initially used in participant classrooms. For example, those who used the conceptual questions to stimulate discussion believe that in the future they will use them as components of exams or homework: "I think I had focused just primarily on using [the material] to stimulate discussion but it is certainly possible to have them as a part of an exam or as a part of a homework."

Alternatively, participants who tried the material on the exams, considered using them as in-class activities in the future: "I can see use it as kind of carrot to get people to, so perhaps put the quiz on Monday morning to make sure that we've read or done the homework beforehand . . . maybe I use it as a group work, I divide them up to groups of 3 or 4 and they talk it out and report back especially if I think it will be something controversial."

Based on the demonstration of experience, some of the participants had detailed insight on how they would use the material in the future. One participant summarized future use of the material in the following way:

"The biggest change that I would make is trying to introduce them sporadically throughout the course of an individual quarter and I would try in touch on more than just two content areas and I think ultimately my preferred implementation medium would be as a technic of formative feedback. I think I would like to introduce a couple of these questions in individual lectures throughout the quarter and different content spaces and then perhaps also include a few of these questions on homework assignments and then in the evaluative exam as well. So I would like to use them both in class and out of class and as a means of assessment and evaluation. I think that would be an optimal implementation."

Potential Barriers to Adoption

To promote the adoption of conceptual questions across the transportation education community, it is critical to identify and ultimately address any potential barriers to adoption. Through the demonstration of the traffic signal conceptual questions, several potential barriers were identified, including the question types (5/9), degree of complexity (3/9), fit within existing course material (3/9), and differences in terminology (1/9).

Issues with the types of questions were the most commonly cited potential barrier. Participants responded that: "the open-ended ones are hard to grade," "ranking . . . is a little bit odd," and "I think there is always a struggle with students doing things like this in a class . . . there are still some students who are stuck in lecture mode that want instructor to tell them what they should learn."

The investment of time in grading student writing is not trivial and likely factors into the adoption of materials. Additionally, the novelty of the ranking task and its requirement for the active participation of students, especially when implemented as a means of formative assessment, could be a source of pushback from students if expectations are not established by the instructor at the outset of their use.

Discussion

This paper describes a holistic effort to develop and test a large set of conceptual exercises and is unique in this broad presentation. An important issue that is investigated but not explored in depth by this work is the relation between problem correctness and students' confidence they got the correct answer. This research proposes five combinations of problems in terms of problem correctness and self-efficacy. However, the implications to learning of these five types remain unexplored. For example, what rationale do students use to justify their confidence? How hard is it to facilitate changing students' confidence when they believe they are correct, but they are not? Is it in fact true that the most challenging misconceptions are those where students are wrong but they are confident that they are right? Or is it possible that questions that students get wrong and are not confident about their answer are the most challenging to correct. More research is needed on the interplay of problem logic and self-efficacy. The interview work cited earlier in this paper focused only on answer and logic, but questions about confidence may provide more insight on how logic and confidence work together on conceptual problems. Insights from research like this could guide teachers to facilitate students overcoming their misconceptions.

The process by which novel curricula are broadly adopted is of critical importance. This research address this concern by actively disseminating the conceptual questions and understanding the perspectives of the instructors who utilized them in their classrooms. Furthermore, the questions have a relatively low activation energy for faculty to implement in the classroom. A single conceptual exercise can be incorporated in as little as 5 min in a lecture, or easily added to a homework or exam. Although this research investigated faculty feedback from implementation of the exercises, it did not investigate how researchers could modify or create new exercises to facilitate overcoming faculty concerns. Such an effort would need to be iterative and long term, but would provide a broader understanding of the kinds of exercises that large percentages of faculty would use, and how exercises could be modified to make them more adoptable. Such intensive and extended studies are rare, but such studies are a natural next step from present work, and can provide the rich and detailed insights from faculty that determine broad adoption.

Conclusions

In summary, this research provides an overview of conceptual question development and insights from student use and faculty implementation. Five distinct patterns were identified, in student answers along with reported self-efficacy: all right, all confident, and all unsure responses, as well as cases where right is confident but wrong lacks confidence and vise a versa. Although some examples are demonstrated for each type of responses, the authors suggest that more work is needed in relation to how students justify their confidence and how this relates to their logic.

Additionally, this paper analyzed instructors' experience using the developed material through semistructured clinical interviews. It was found that conceptual questions dealing with fundamental aspects of traffic engineering, such as cycle length, red clearance, and yellow change intervals, were most frequently selected by instructors. Multiple choice questions were the most popular type of questions and that enriching exam materials and creating challenging discussions were the primary goals of implementation.

More in-depth studies are needed with faculty to understand how more conceptual exercises across multiple disciplines can be developed and adopted. These two works could come together in the development of exercises that explore both logic and confidence, with the goal of addressing both explicitly with the exercises. Development of such adoptable exercises would be impactful to transportation engineering, but also are needed across STEM disciplines.

Acknowledgments

This material is based in part upon work supported by the National Science Foundation under Grant No. DUE-1140400. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Bandura, A. (1997). *Self-efficacy: The exercise of control*, W.H. Freeman and Company, New York.
Black, P., and Wiliam, D. (1998). "Assessment and classroom learning." *Assess. Educ. Principles Policy Pract.*, 5(1), 7–74.
Brown, S., and Poor, C. (2010). *Ranking tasks for mechanics of materials*, Prentice Hall, Englewood Cliffs, NJ.
Chi, M. T. H. (2009). "Active-constructive-interactive: A conceptual framework for differentiating learning activities." *Top. Cognit. Sci.*, 1(1), 73–105.
Chung, G. K., Shel, T., and Kaiser, W. J. (2006). "An exploratory study of a novel online formative assessment and instructional tool to promote students' circuit problem solving." *J. Technol. Learn. Assess.*, 5(6), 1–27.
Dedoose version 6.0.24 [Computer software]. <http://www.dedoose.com/>.
Garber, N., and Hoel, L. (2015). *Traffic and highway engineering*, 5th Ed., Cenage Learning, Stamford, CT.
Hurwitz, D., Brown, S., Islam, M., Daratha, K., and Kyte, M. (2014). "Traffic signal system misconceptions across three cohorts: Novice

students, expert students, and practicing engineers." *Transp. Res. Rec.*, 2414, 52–62.
Ludvigsen, K., Krumsvik, R., and Furnes, B. (2015). "Creating formative feedback spaces in large lectures." *Comput. Educ.*, 88, 48–63.
Mannering, F., and Washburn, S. (2013). *Principles of highway engineering and traffic analysis*, 5th Ed., Wiley, Hoboken, NJ.
NRC. (1999a). *Bridging research and practice*, National Academy Press, Washington, DC.
NRC. (1999b). *How people learn: Brain, mind, experience, and school*, National Academy Press, Washington, DC.
O'Kuma, T. L., Maloney, D. P., and Hieggelke, C. J. (2003). *Ranking tasks in physics: Student edition*, Benjamin Cummings, San Francisco.
Roess, R., Prassas, E., and McShane, W. (2011). *Traffic engineering*, 4th Ed., Prentice Hall, Englewood Cliffs, NJ.
Rogers, E. M. (1995). *Diffusion of innovations*, 14th Ed., Free Press, New York.
R version 3.2.0 [Computer software]. The R Foundation for Statistical Computing.
Yan, Z., and Cheng, E. C. K. (2015). "Primary teachers' attitudes, intentions and practices regarding formative assessment." *Teach. Teach. Educ.*, 45, 128–136.