International Journal of Engineering Education Vol. 33, No. 1(A), pp. 225–235, 2017
Printed in Great Britain
0949-149X/91 $3.00+0.00

Evaluating Students' Understanding of Statics Concepts Using Eye Gaze Data

YOUYI BI and TAHIRA REID
School of Mechanical Engineering, Purdue University, 585 Purdue Mall, West Lafayette, IN 47907, USA. E-mail: bi9@purdue.edu

In engineering courses, exams and homework assignments are among the standard tools used to assess students' performance and comprehension of course material. However, they do not always provide opportunities to reveal whether students truly understand related engineering concepts. This paper seeks to bridge that research gap by using eye-tracking technology to observe how students solve statics problems. In a within-subject experiment, twenty participants were asked to solve nine statics problems shown on a computer display. A non-invasive eye-tracker was used to record participants' eye movements during the problem solving process. Participants were then asked to explain how they solved three representative problems. The results show that different eye gaze patterns exist between those who solved problems correctly and those who solved them incorrectly. For the specific concepts involved in solving these problems, those who correctly understood the concepts also exhibited different eye gaze patterns than those who did not. We also found that students' spatial visualization skills positively correlate with their performance when solving statics problems. This investigation showed that eye gaze data has the potential to serve as a diagnostic tool to discern how students solve statics problems and understand related engineering concepts. These results may provide insight into students' problem-solving strategies and difficulties, and help instructors choose more adaptive teaching methods for students.

Keywords: statics; problem-solving; concept inventory; eye-tracking

1. Introduction

Engineering is a challenging discipline of study. According to Budny's research on freshman performance in engineering courses at Purdue University [1], over a span of fifteen years (1978–1993) approximately 36% of the entering engineering students failed to complete the freshman requirements and thus did not transfer into one of the professional schools of engineering. Interviews with these students indicated that the main reason for leaving engineering was difficulty with calculus, chemistry, or physics. Instructors have developed multiple assessment tools, such as homework, projects, and exams to assess students' grasp of subject matter and their ability to apply new concepts [2].

These assessment tools, however, do not always encourage a deep understanding of subject matter [3]. For instance, Cooper et al. [4] argued that in chemical engineering courses, examinations and quizzes—including both short answer and multiple choice—give very little insight into the problem-solving process itself. Sometimes it is not easy for teachers and evaluators to determine whether a student's unsatisfactory performance is the result of misunderstanding the topics discussed, weakness in spatial thinking, or other factors [5]. Students who choose the correct answers in these tests may not completely understand the related concepts. Many investigators suggest think-aloud protocols and oral tests (i.e., asking students to introduce their approach to solving problems verbally) as the better way to examine whether students truly understand the critical concepts and provide a more nuanced picture of the problem-solving process [4, 6–8]. However, universities rarely administer oral tests due to prohibitive time constraints.

Fortunately, eye-tracking technology can provide a way for researchers to observe how students solve engineering problems through their eye movements. Eye-tracking research is based on the "eye-mind" hypothesis which states that people look at what they are thinking about [9]. It assumes that people fixate on a specific area of a problem diagram longer when they encounter difficulties or are confused [10]. Utilizing eye-tracking technology allows us to observe the visual attention patterns of students while solving engineering problems and evaluate their understanding of specific concepts. Once such patterns of visual attention are discovered, they can be leveraged to determine which concepts are the most challenging, and may enable instructors to provide students with more targeted help.

In this paper, we present the results of a study that examined the visual attention of 20 undergraduate students while they solved statics problems. The students were asked to solve nine statics problems displayed on a computer screen with their eye movements recorded. Students were classified into "correct solvers" and "incorrect solvers" for each problem. Correct solvers chose the correct answer to the problem while incorrect

* Accepted 25 October 2016.

solvers did not. They were then asked to explain how they solved three of the nine problems verbally. Based on their explanations, students were categorized as two groups: Group 1 "correctly understood" and Group 2 "incorrectly understood" for each tested statics concept. Students in the correctly understood group were able to explicitly explain the key concept involved in solving the problem while students in the "incorrectly understood" group were not. In the remainder of this paper, Section 2 provides background on problem-solving, visual perception and eye tracking methods and proposes two hypotheses. The experimental methods and results are presented in Sections 3 and 4, respectively. Section 5 provides detailed explanations for the study results and conclusions are given in Section 6.

2. Related literature

2.1 Problem-solving and visual attention patterns

In engineering education research, it has been appreciated that effective assessment is important for monitoring students' progress and providing feedback to students [11]. Traditional assessment tools such as homework and exams are broadly used though certain limitations exist. For example, Boud [12] pointed out that students who perform well on university examinations retain fundamental misconceptions about key concepts in the subjects they have passed. Mooney et al. [13] argued that multiple choice questions assess only recall and recognition, and promote a surface approach to studying. This view is shared widely [14–18].

Another limitation of these methods is that they do not always provide insights on students' problem solving process (e.g., multiple choice questions). Visual attention patterns, or how individuals perceive and interpret problem diagrams, have been used in problem-solving research to evaluate their profession-related expertise. Grant et al. [19] studied how eye movements reveal critical aspects in solving a tumor-and-lasers radiation problem, a classical insight problem developed by Duncker in 1945 [20]. In one experiment, they found that certain fixation patterns correlate with success in solving this particular problem. In a second experiment, they found that perceptually highlighting the critical diagram component, identified in the first experiment, significantly increased the frequency of correct solutions. Hegarty [5] also pointed out that processes for manipulating spatial information are also central to mechanical problem solving. Similar observations can be found in other contexts, such as the examination of chest X-ray (CXR) films [21], in aviation [22], during driving [23], solving anagram problems [24], and when identifying explosives at airports [25].

Statics problems are well-defined and focus on the ability to understand and interpret mechanics diagrams. Individuals solving such problems must be able to determine the important visual cues in the diagram and apply the correct concepts during the problem solving task. Thus, the visual attention patterns on the problem diagrams and answers can be used to indicate students' levels of understanding and their ability to apply related statics concepts.

2.2 The application of eye-tracking in problem-solving research

Problem solving involves information processing and decision making, which can be monitored by various process tracing methods. Example methods include surveys/self-report [26], computer-based information board paradigms (e.g. Mouselab [27]) and think-aloud protocols [28]. However, these techniques sometimes influence decision behavior [29] and might hinder participants' automatic processing in information search [30]. In contrast, biometric signals, such as eye gaze data, provide tracing information without hindering automatic information acquisition processes. These signals are promising to provide insights on understanding individuals' problem-solving patterns [31].

According to Just and Carpenter's eye-mind hypothesis [9], people fixate longer on a specific problem diagram when they are attracted or confused. The main metrics used in eye-tracking include: (1) fixations: eye movements that stabilize the retina over a stationary object of interest (in this study, a fixation occurs when the eyes focus on a specific area for more than 100 milliseconds); (2) fixation time: a measure of the duration of the fixation on a specific area; and (3) average fixation duration: the duration time per fixation [32]. The location and duration of fixations is directly related to the locus and difficulty of cognitive processing [33]. Thus, tracking eye movements may provide insight into what visual information is being processed and how difficult this information is to process, which may serve as an additional measure for people's thinking processes [34]. This technology has been applied in many areas for visual attention research, including attention neuroscience [35], reading [36], visual inspection [37], studying worked-out mechanics problems [38], and arithmetic problem solving [39, 40].

Based on the theories explaining the reproducibility of expert superiority in visual domains [41], high performers and low performers will show different eye gaze patterns (fixation counts, fixation time, etc.) while solving problems with visual elements. Madsen found that while solving physics problems, correct solvers spent more time attending to relevant areas, whereas incorrect solvers spent more time looking at novice-like areas [42]. Consistent results can be found in Tsai et al.'s research on visual attention for solving multiple-choice science problems [43]. These studies show the potential of applying eye-tracking in problem solving, which provides researchers with more comprehensive ways to understand individuals' problem-solving behavior. Thus, in this study we hypothesize that performers of different expertise levels will show different eye gaze patterns while solving statics problems on a computer screen.

2.3 Summary

The existing literature has shown the utilization of individuals' visual perception in problem-solving research. The literature has also emphasized the value of eye gaze data in obtaining visual attention patterns during the problem-solving process and in facilitating people's problem solving performance. However, there is limited work evaluating how the eye gaze data reveal students' understanding related concepts in solving engineering problems. This paper seeks to bridge that research gap by testing the following hypotheses:

H1: Different visual attention patterns exist in correct and incorrect solvers when solving statics problems.
H2: Different visual attention patterns exist in those who correctly and incorrectly understand the related statics concepts when solving statics problems.

3. Methods

To test the two hypotheses given in Section 2, we designed a study including three parts. Part 1 consists of a computer-based survey that presented 9 statics problems to participants while their eye movements were recorded by an eye-tracker. In Part 2, participants verbally explained how they solved three of the nine problems which was video recorded. The three problems selected were based on prior work that showed these particular problems were well designed for students with appropriate difficulty and discrimination levels [44]. The experimenter can determine if the participants truly understood the key concepts involved in solving each problem by watching the replay of these videos. In Part 3, participants finished a set of tests about their background information which included a spatial thinking ability test, a questionnaire on their learning styles and a demographics survey. In this section the test materials, subject demographics, and experiment procedures are described.

3.1 Test materials

The statics problems come from the Concept Assessment Tool for Statics (CATS) [45], developed by Prof. Paul Steif at Carnegie Mellon University. CATS is a multiple choice test that assesses students' conceptual knowledge of statics. The test consists of 27 questions that capture 9 distinct concepts and include distractors (wrong answers) that have been constructed based on observations of student work. Its reliability and quality have been confirmed based on tests administered to more than one thousand students and ten classes in seven US universities [46]. Due to the time limit of the eye-tracking experiment, we selected 9 CATS items covering 3 statics topics for our study (3 items for each topic). Table 1 lists the topics covered, key concepts involved and difficulty of each item. For example, CATS items 25, 26 and 27 cover the topic "equilibrium" in engineering statics. To solve these three problems correctly, students need to understand the two key concepts involved: (1) consider the balance of all forces, (2) consider the balance of all moments. The difficulty values indicate the proportion of correct problem solvers to the previously tested population for each problem [46]. Thus, the CATS items with larger difficulty values

Table 1. The CATS items used in the study

Topics covered	Key concepts involved	CATS Item No.	Difficulty [47]
Pin and Slot: direction of force between the roller and the rolled surface	1. The reaction force must be perpendicular to the contact surface. 2. No moment would exist at the contact surface.	**13** 14 15	0.697 0.696 0.735
Negligible Friction: direction of force between frictionless bodies in point contact	1. The reaction force must be perpendicular to the tangent of the contact surface. 2. No moment would exist at the contact point.	16 18 **17**	0.281 0.576 0.264
Equilibrium: consideration of both force and moment balance in equilibrium	1. Consider the balance of all forces at all directions in equilibrium. 2. Consider the balance of all moments in equilibrium	25 26 **27**	0.615 0.161 0.487

would be easier to students. Students were asked to explain how they solved three problems (CATS items 13, 17 and 27) verbally in Part 2 of the study based on insights from prior work on these 3 problems [44].

Fig. 1 shows one CATS item presented to participants on a computer display (CATS No. 25). For this problem, participants are tested on the "Equilibrium" topic, in which they must consider both the force and moment balance. The dashed rectangles indicate the Areas of Interest (AOIs) for this visual stimulus, including the *problem statement* area (in the top), the *problem diagram* area (in the bottom left) and five options (in the bottom right). Within the five options, we are interested in two kinds of answers: *correct answer* and *indicator answer*. For example, in the problem presented in the Fig. 1, the correct answer is (d), because both the rightward force and the moment should be applied to end A in order to balance the whole member. If participants correctly understood the first concept involved in the "Equilibrium" topic (see Table 1), they would eliminate answer (a) quickly, as the direction of the applied force is different than the given leftward force. Also, compared to answer (e), the reactions involved in answer (a) are much simpler and easier for participants to judge. Thus, answer (a) is an indicator answer for evaluating participants' understanding of this concept. These AOIs are created for analyzing participants' visual attention patterns during the problem-solving process. Participants did not see the rectangles during the test.

The revised Purdue Spatial Visualization Test: Rotation (PSVT: R) was used to test participants' spatial thinking ability. This test was initially developed by Guay, 1976 [48] and then modified by Yoon, 2011 [49]. It consists of a total of 30 questions that must be completed in 20 minutes to completely assess students' spatial visualization abilities. PVST: R is considered one of the best measures of an individual's spatial ability [50]. The Index of Learning Styles questionnaire (ILS) [51] was used to assess participants' learning styles based on the Felder-Silverman model. This model is reliable, valid and suitable for reflecting engineering students' learning style profiles [52, 53], and providing an indication of students' possible strengths, tendencies or habits in learning.

3.2 Participants

Institutional Review Board approval was gained from the Purdue IRB before conducting the experiment. A total of 20 engineering undergraduate students from Purdue University participated in this study, ranging in age between 18 and 24 years. Seventeen were male and 3 were female. Fifteen of them were sophomores, 2 were juniors and 3 were seniors. Their majors included Industrial Engineering, Biomedical Engineering and Mechanical Engineering. All, had already taken a statics course. Ten participants obtained a grade of "A–" or above, 9 got "B–, B or B+" and one got "C–, C or C+".

Participants were recruited by email. They also had to meet the inclusion criteria suggested by

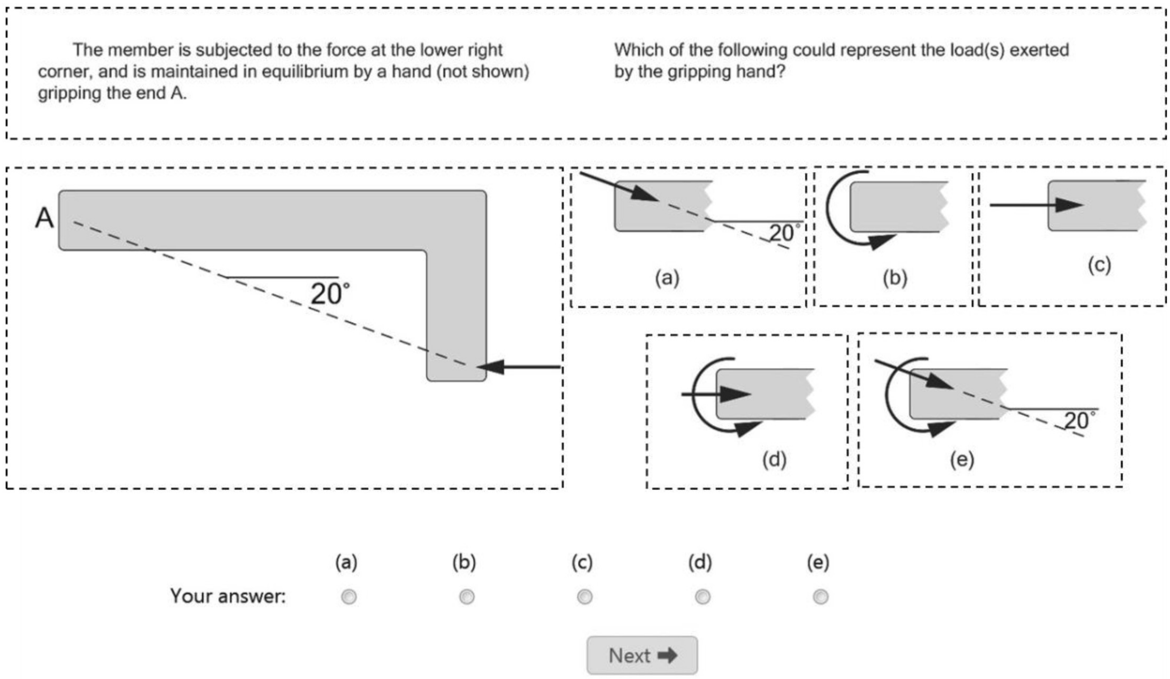


Fig. 1. A CATS item used in the study (reprinted with permission of the author of the CATS [45]) and its Areas of Interest (AOIs).

Pernice and Nielsen [54] to satisfy the experimental conditions of eye tracking research:

- Have normal to corrected vision (contact lenses and glasses are okay, except for bifocals, trifocals, layered lenses or regression lenses).
- Do not have glaucoma, cataracts, eye implants, or permanently dilated pupils.
- Can read a computer screen and the Web without difficulty.
- Do not need a screen reader, screen magnifier or other assistive technology to use the computer and the Web.

3.3 Experimental procedure

The participants completed the study individually. After passing pre-screening, qualified participants were introduced to the purpose and procedures of the study by the experimenter. Then participants were required to sit in front of a computer display and adjust their sitting positions to ensure the successful calibration of the eye-tracker (Tobii X-60, Tobii Technology AB, Danderyd, Sweden). The eye-tracker has a sampling rate of 60 Hz. If the calibration result was poor, participants had to do a recalibration. If the calibration result was good, participants would see an introduction screen and proceed to the test. Each participant saw the nine statics problems in a randomized order.

After finishing Part 1, participants were asked to explain how they solved three of nine problems with pen and paper. Their explanations were recorded with a video camera. Typical questions asked by the experimenter include: "Please tell me how you solved this problem.", "Please tell me why you selected your response.", "Why do you think your selected response is correct?", "Can you tell me why other responses are incorrect?" The experimenter then determined whether the participant truly understood the involved statics concepts (see Table 1) based on their explanation and divided them into two groups: correctly understood and incorrectly understood for each concept.

After finishing Part 2, participants were asked to complete the PSVT: R test (with a time limit of 20 minutes), an ILS questionnaire and a survey on their background information. The whole process of the study typically took 50 minutes. Participants were compensated 10 dollars at the end of the study for their participation.

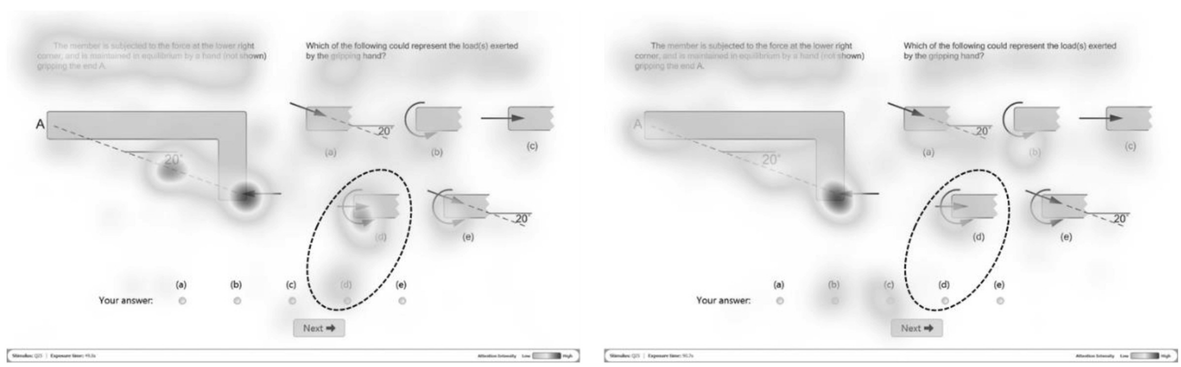
4. Results

In this section, participants' performance in solving statics problems associated with their eye gaze patterns are presented. The iMotions software (iMotions, Inc., Cambridge MA) is used for analyzing eye gaze data. The correlations between participants' statics problem-solving performance and other factors (spatial thinking ability, learning style and background information) are also provided.

4.1 Eye gaze patterns of correct and incorrect statics problem solvers

Fig. 2 illustrates the comparison of the eye gaze heatmaps for one statics problem (CATS item 25) between correct solvers and incorrect solvers. The deeper color indicates the areas receiving more eye fixations. This figure intuitively shows that performers of different level of success showed different eye gaze patterns: correct solvers spent more time in the correct answer (d) than incorrect solvers did (notice the areas circled by the dashed eclipses).

Fig. 3 presents the comparison of the eye gaze patterns between correct and incorrect statics problem solvers in three AOIs (problem statement, problem diagram and correct answer). This comparison combines the data across all 9 statics problems.



(a) (b)

Fig. 2. The comparison of the eye gaze heatmaps for CATS Item 25 between (a) correct solvers and (b) incorrect solvers (see [55] for the polychromatic version of the heatmaps).

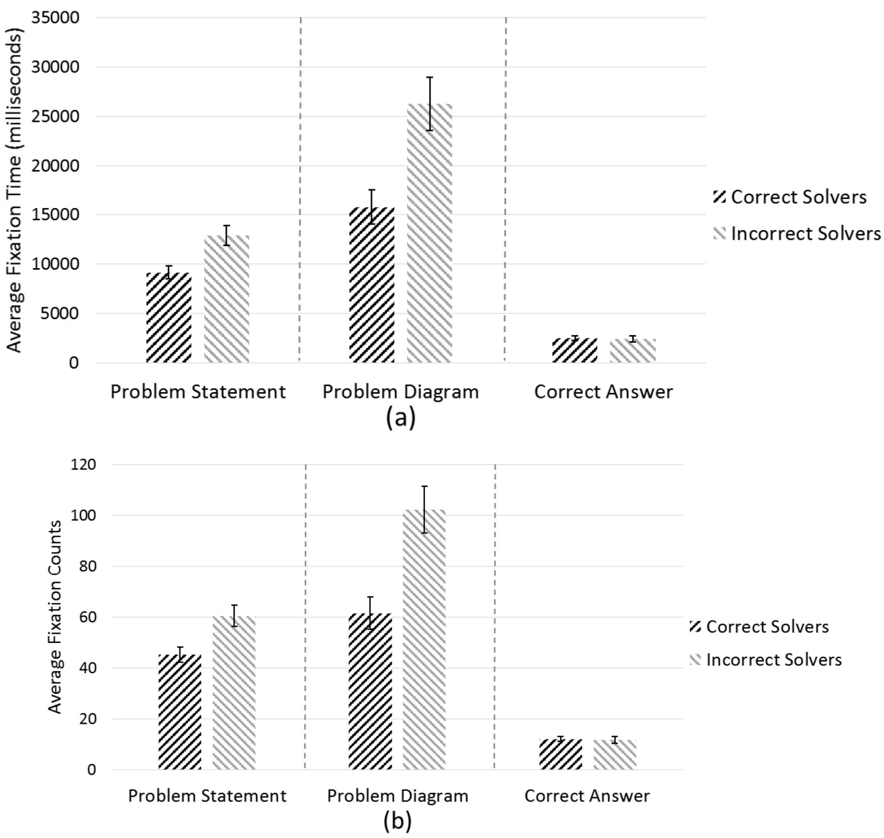


Fig. 3. The comparison of the eye gaze patterns between correct and incorrect solvers. (a) Average Fixation Time; (b) Average Fixation Counts. Bars indicate ± the standard error.

Notice that after combining the data across all 9 statics problems, eye gaze metrics (the average fixation time, and average fixation counts) do not differ as significantly in the correct answer area. However, we can observe the differences in the areas of problem statement and problem diagram. The Welch two sample t-test ($\alpha = 0.05$) was used to test the differences. As shown in Fig.3(a), we found that correct solvers had less fixation time on the area of problem statements than incorrect solvers ($t(149) = -3.18$, $p = 0.001$), and correct solvers also had less fixation time on the area of problem diagrams ($t(149) = -3.24$, $p = 0.001$). Similarly, in Fig. 3(b), correct solvers had less fixation counts on the area of problem statements than incorrect solvers ($t(149) = -3.00$, $p = 0.002$), and they also had less fixation counts on the area of problem diagrams ($t(149) = -3.63$, $p < 0.001$). These results indicate that correct and incorrect statics problem-solvers indeed showed different eye gaze patterns in perceiving the problem statement and diagram during the process of solving the statics problems, which is consistent with our hypothesis in Section 2.2.

4.2 Eye gaze patterns of "correctly and incorrectly understood" participants

In Part 2 of the study, participants were asked to explain how they solved three representative problems (CATS items 13, 17 and 27) verbally. They were then classified into two groups: correctly understood and incorrectly understood for each concept based on their explanation of the problem-solving process. This classification work was implemented by two PhD candidates in Mechanical Engineering independently, who are quite familiar with statics. The Cohen's Kappa of their rating results is 0.81, which indicates that the interrater reliability in the classification is acceptable. Fig.4 presents the comparison of the eye gaze patterns between correctly understood group and incorrectly understood group in four AOIs (problem statement, problem diagram, correct answer and indicator answer). This comparison combines the data across participants' understanding of the 6 concepts involved in solving the 3 statics problems (2 concepts for each problem).

As shown in Fig.4 (a), the correctly understood group had less fixation time on the areas of problem statements than the incorrectly understood group ($t(80) = -3.09$, $p = 0.001$); the same is true for the area of problem diagrams ($t(69) = -1.89$, $p = 0.031$). This trend is consistent for the fixation time on the areas of correct answer ($t(83) = -0.82$, $p = 0.208$) and indicator answer ($t(109) = -1.63$, $p = 0.053$), though

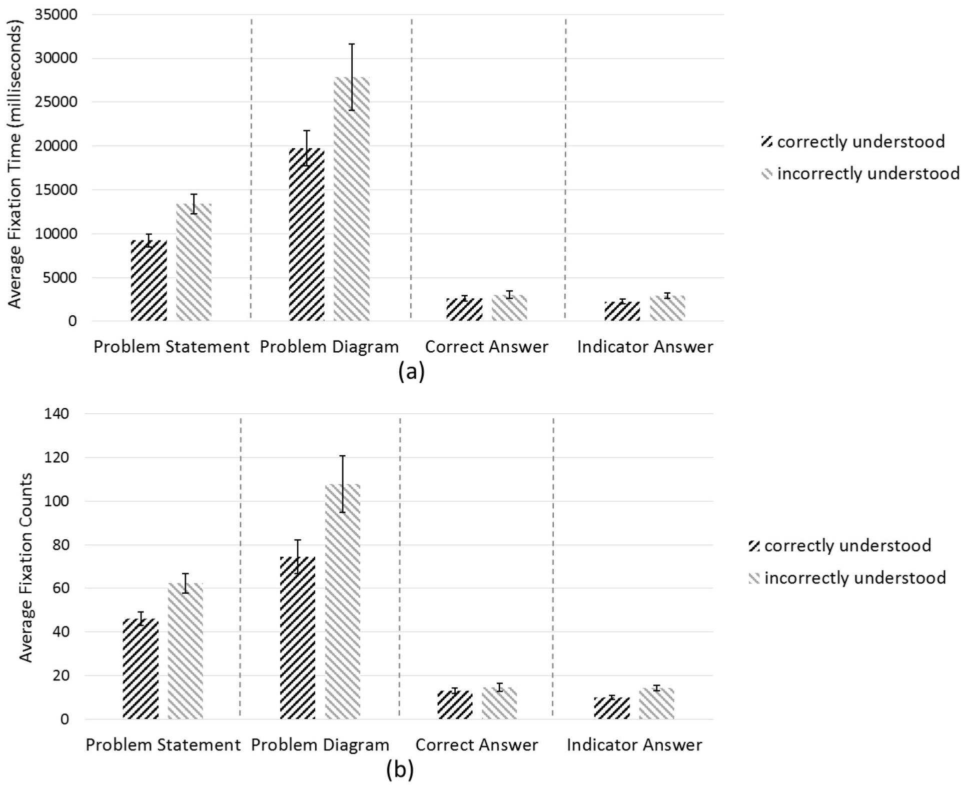


Fig. 4. The comparison of the eye gaze patterns between correctly understood group and incorrectly understood group. (a) Average Fixation Time; (b) Average Fixation Counts. Bars indicate ± the standard error.

the differences are not significant. In Fig. 4(b), the correctly understood group had significantly less fixations than the incorrectly understood group on the areas of problem statements ($t(87) = -3.04$, $p = 0.002$), problem diagrams ($t(75) = -2.26$, $p = 0.013$) and indicator answer ($t(88) = -2.63$, $p = 0.005$). A similar trend can be observed for the fixation time on the areas of correct answer ($t(84) = -0.66$, $p = 0.254$), though the differences are not significant. These results indicate that the correctly understood group and incorrectly understood group showed different eye gaze patterns in perceiving the problem statement and diagram during the process of solving the statics problems. Also, the correctly understood group fixated on the indicator answer less frequently than incorrectly understood group.

4.3 *Correlation between participants' statics problem-solving performance and other factors*

In Fig. 5 and Fig. 6, participants' statics problem-solving performance (CATS performance) are plotted against their spatial thinking ability (PSVT: R performance) and average problem-solving time, respectively. Each dot represents a participant (labeled by the subject number). The CATS performance ranges from 0 to 9, and the

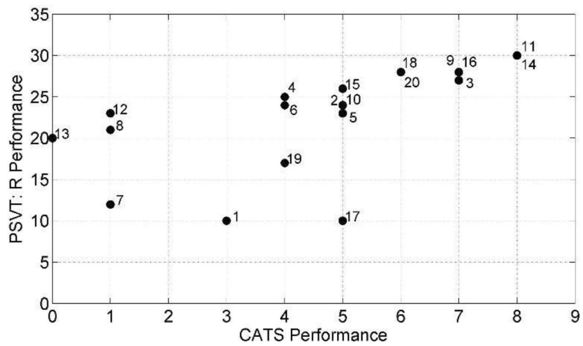


Fig. 5. The correlation between participants' CATS performance and PSVT: R performance.

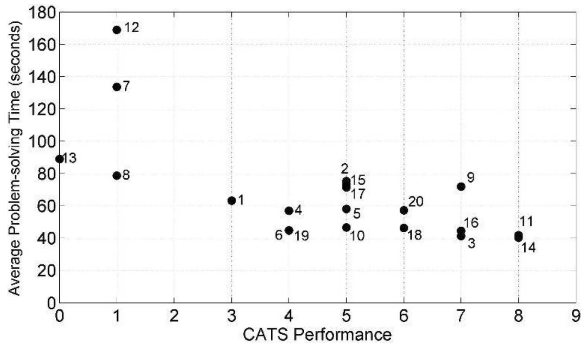


Fig. 6. The correlation between participants' CATS performance and average problem-solving time.

value indicates how many statics problems the participants solved correctly. Analogously, the PSVT: R performance ranges from 0 to 30 indicating the number of correctly answered PSVT: R questions. It is interesting to find that participants' CATS performance correlates with their PSVT: R performance positively (r = 0.618, p = 0.004), whereas it correlates with their average problem-solving time negatively (r = –0.7, p = 0.001). We did not find significant correlations between participants' CATS performance and other factors (i.e. learning styles, statics course grades and mechanics-related experience).

5. Discussion

Consistent with the previous findings of problem-solving research in other domains, we found that, when solving statics problems, correct solvers had less fixation time and less fixation counts on the area of problem statements and problem diagrams than incorrect solvers. This result indicates that the hypothesis H1 cannot be rejected: correct and incorrect statics problem-solvers showed different eye gaze patterns during the process of solving the statics problems. Hegarty [5] used to argue that when solving mechanical problems, experts are faster and expend less effort than novices, because once experts select a problem schema, they automatically access the procedures for solving the problem and these procedures can be carried out without further effort. Furthermore, the protocols of novices contain more meta-statements expressing sub-goals and uncertainties. Considering the situation of solving statics problems, we can also infer that correct solvers were faster to extract the key information from the problem statements and diagrams and apply the correct problem-solving protocols than incorrect solvers. However, we didn't find significant differences in the fixation time and fixation counts on the areas of correct answer. This observation implies that incorrect solvers spent nearly the same amount of visual attention on the correct answer as correct solvers did, however, they still failed to accept it as the final choice after their mental reasoning process.

It is interesting to observe that participants who correctly and incorrectly understood related statics concepts also show different eye gaze patterns, which indicates that the hypothesis H2 cannot be rejected. Our experimental results show that correctly understood group spent less fixation time and had less fixation counts on the areas of problem statements and diagrams than incorrectly understood group. Additionally, correctly understood group had significantly less fixation counts on the area of indicator answer than incorrectly understood group. Similar trends can be seen with respect to the metric of fixation time, though the difference is not significant (p = 0.053). This result implies that the correctly understood group was able to eliminate obviously wrong answers more quickly than the incorrectly understood group.

However, there was not always an exact correlation between students' problem-solving and concept-understanding. For example, though subject 4, 9 and 19 chose the correct answer when solving CATS item 13, they all believed that there would be a moment existing in the contact point of the frictionless pin-slot joint during the interview, indicating that they didn't understand the second involved concept. In contrast, although subject 10 and 19 chose the wrong answer when solving CATS item 17, their explanation of the problem-solving process indicated that they truly understood the concepts despite selecting the wrong answer. These observations suggest that students' answers in multiple choice questions may not always appropriately reflect their true level of understanding of relevant concepts. Besides oral tests (such as the interviews in our study), eye gaze data could also be utilized as an auxiliary indicator based on the findings in this paper.

Correlations between participants' CATS performance and their background information provide us with additional insights on the factors that influence students' statics problem solving. Our results show that participants' PSVT: R performance correlates positively with their statics problem solving performance. This finding suggests that spatial thinking ability plays an important role in solving statics problems, since participants had to visualize the free body diagrams in their minds without using paper and pen. We also found that participants' CATS performance correlates negatively with their average problem solving time, which is consistent with the eye gaze results discussed above. However, large variations exist within the participants who possessed the same CATS performance or similar average problem-solving time. For example, subject 12 and 8 both only answered one CATS problem correctly, but their average problem solving time showed large differences (see Fig. 6). This may imply that some students were more careful and meticulous in their problem solving. On the other hand, subject 6 and 11 had approximately equivalent average problem-solving time but their CATS performance was dramatically different. This suggests that average problem-solving time would not be a perfect indicator of students' problem-solving performance. Furthermore, we did not find any significant correlations between participants' CATS performance and their learning styles, statics course grades, or mechanics-related experience. This suggests that unique learning styles would not greatly impact individuals' problem solving performance, and each student can develop his or her own ways to learn and understand concepts well.

6. Conclusions

The present study verified that participants with different levels of success showed different eye gaze patterns when solving statics problems. Correct problem solvers had less fixation time and less fixation counts on the area of problem statements and problem diagrams than incorrect solvers. Participants who correctly understood related concepts showed the same trend of eye gaze patterns. Additionally, the correctly understood group had significantly less fixation counts on the area of the indicator answer than the incorrectly understood group. We also found that participants' spatial thinking ability correlate with their statics problem solving performance positively. Participants' learning styles and mechanics-related experience were not found to have an influence on their statics problem solving performance. These results suggest that eye gaze data has the potential to serve as a diagnostic tool to evaluate how well students understand engineering concepts during solving statics problems. This research may also help instructors choose more adaptive teaching strategies for students with different levels of understanding of engineering concepts.

A limitation of this study is the fact that the incentive for participation was not performance-based, and therefore participants may not have put forth their best effort. Other research opportunities include exploring whether students' problem solving performance will be improved by enhancing their spatial thinking abilities along with their understanding of key concepts in mechanics. Researchers have found that hands-on experience [56], peer-to-peer collaboration, assessment and feedback [57–59] in engineering classes can have an influence on students' understanding of engineering concepts. Future work can combine the use of eye-tracking as a diagnostic tool with some of these in-class activities to help remedy students' learning challenges. As Lai et al. [60] suggested, "interactive learning systems embedded with eye-tracking equipment may dynamically diagnose students' learning states and needs as well as provide instant help or adapted scaffolding materials according to the eye movement data tracked by the systems".

Although eye-tracking has not been used in engineering education widely, our research findings have the potential to be applied in other engineering courses as well. This is due to the close connection of the problem-solving processes of mechanical engineering courses. For example, Fang and Lu [61] found that a student's performance in Statics and cumulative GPA play the two most significant roles in governing the student's performance in Dynamics. On the other hand, science educators have used eye-tracking to study students' performance. For example, Chen et al. [62] found that students' eye movement behavior can successfully predict their computer-based assessment performance of physics problems. Their results show that students who provided correct answers had shorter saccade behaviors than those who provided incorrect answers to picture presentations, which are comparable to our results. Future research can be extended to use eye-tracking in other engineering courses, which enables researchers to further understand the consistencies and inconsistencies of students' eye gaze patterns across solving problems of various engineering courses.

Eye-trackers are becoming more affordable where some low-cost solutions are emerging [63–66]. Thus, the lower costs may enable researchers and universities to acquire a system for educational studies. However, few studies have been published using these low-cost solutions, and no studies have presented and compared results from both these solutions and those of commercial eye-trackers. Therefore, researchers will need to characterize these systems by replicating the findings of others. In addition, further work is needed to determine if "eye training" should be considered as a teaching/learning objective in engineering courses, since this practice has been extensively studied in basic cognitive skills and clinical medicine [67, 68]. Furthermore, running eye-tracking studies can be time-intensive. Therefore, future work can also be extended to include the use of existing eye gaze data to develop prediction models of eye gaze patterns in problem solving.

Acknowledgements—The authors gratefully acknowledge the support, help and advice from Dr. So Yoon Yoon, Prof. Chuck Krousgrill, Prof. Ruth Streveler, Prof. Paul Steif, Dr. Dana Denick, Rasmine Sabra and the members of REID lab.

References

1. D. Budny, W. LeBold and G. Bjedov, Assessment of the Impact of Freshman Engineering Courses, *Journal of Engineering Education*, **87**(4), 1998, pp. 405–411.
2. M. Richards-Babb, J. Drelick, Z. Henry and J. Robertson-Honecker, Online Homework, Help or Hindrance? What Students Think and How They Perform, *Journal of College Science Teaching*, **40**(4), 2011, pp. 81–93.
3. D. J. Hargreaves, Student Learning and Assessment Are Inextricably Linked, *European Journal of Engineering Education*, **22**(4), 1997, pp. 401–409.
4. M. M. Cooper, C. T. Cox Jr., M. Nammouz, E. Case and R. Stevens, An Assessment of the Effect of Collaborative

Groups on Students' Problem-solving Strategies and Abilities, *Journal of Chemical Education*, **85**(6), 2008, pp. 866–872.
5. M. Hegarty, Knowledge and Processes in Mechanical Problem Solving, *Complex problem solving: Principles and Mechanisms*, Psychology Press, New York, NY, 1991, pp. 253–285.
6. K. A. Ericsson and H. A. Simon, Verbal Reports as Data, *Psychological Review*, **87**(3), 1980, pp. 215–251.
7. M. T. Chi, M. Bassok, M. W. Lewis, P. Reimann and R. Glaser, Self-explanations: How Students Study and Use Examples in Learning to Solve Problems, *Cognitive Science*, **13**(2), 1989, pp. 145–182.
8. M. T. Chi, P. J. Feltovich and R. Glaser, Categorization and Representation of Physics Problems by Experts and Novices, *Cognitive Science*, **5**(2), 1981, pp. 121–152.
9. M. A. Just and P. A. Carpenter, Eye Fixations and Cognitive Processes, *Cognitive Psychology*, **8**(4), 1976, pp. 441–480.
10. P. Du and E. F. MacDonald, Eye-Tracking Data Predict Importance of Product Features and Saliency of Size Change, *Journal of Mechanical Design*, 136(8), 2014, pp. 081005-1-081005-14.
11. P. Black and D. William, Assessment and Classroom Learning, *Assessment in Education*, **5**(1), 1998, pp. 7–73.
12. D. Boud, Ensuring that Assessment Contributes to Learning, *Proceedings of the International Conference on Problem Based Learning in Higher Education*, University of Linkoping, Sweden, 1995, pp. 13–20.
13. G. A. Mooney, J. G. Bligh and S. J. Leinster, Some Techniques for Computer-Based Assessment in Medical Education, *Medical Teacher*, **20**(6), 1998, pp. 560–6.
14. N. Falchikov, Chapter 2-What's Wrong with Traditional Assessment, *Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education*. RoutledgeFalmer, New York, NY, 2005, pp. 32–40.
15. D. Wolf, J. Bixby, J. III Glenn and H. Gardner, To Use Their Minds Well: Investigating New Forms of Student Assessment, *Review of Research in Education*, **17**, 1991, pp. 31–73.
16. D. B. Swanson, G. R. Norman and R. L. Linn, Performance-based Assessment: Lesson from The Health Professionals, *Educational Researcher*, **24**(5), 1995, pp. 5–11.
17. K. M. Scouller, The Influence of Assessment Method on Student's Learning Approaches: Multiple Choice Question Examination versus Assignment Essay, *Higher Education*, **35**, 1998, pp. 453–72.
18. P. Davies, Computerized Peer Assessment, *Innovations in Education and Training International*, **37**(4), 2000, pp. 346–55.
19. E. R. Grant and M. J. Spivey, Eye Movements and Problem Solving Guiding Attention Guides Thought, *Psychological Science*, **14**(5), 2003, pp. 462–466.
20. K. Duncker and L. S. Lees, On Problem-solving, *Psychological Monographs*, **58**(5), 1945.
21. S. A. Vitak, J. E. Ingram, A. T. Duchowski, S. Ellis and A. K. Gramopadhye, Gaze-augmented Think-aloud as an Aid to Learning, *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*. ACM, 2012, pp. 2991–3000.
22. W. L. Ottati, J. C. Hickox and J. Richter, Eye Scan Patterns of Experienced and Novice Pilots During Visual Flight Rules (VFR) Navigation, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* SAGE Publications, **43**(1), 1999, pp. 66–70.
23. G. Underwood, P. Chapman, N. Brocklehurst, J. Underwood and D. Crundall, Visual Attention While Driving: Sequences of Eye Fixations Made by Experienced and Novice Drivers, *Ergonomics*, **46**(6), 2003, pp. 629–646.
24. J. J. Ellis, M. G. Glaholt and E. M. Reingold, Eye Movements Reveal Solution Knowledge Prior to Insight, *Consciousness and Cognition*, **20**(3), 2011, pp. 768–776.
25. D. Schuster, J. Rivera, B. C. Sellers, S. M. Fiore and F. Jentsch, Perceptual Training for Visual Search, *Ergonomics*, **56**(7), 2013, pp. 1–15.
26. R. E. Nisbett and T. D. Wilson, Telling More than We Can Know: Verbal Reports on Mental Processes, *Psychological Review*, **84**(3), 1977, pp. 231–259.
27. J. W. Payne, J. R. Bettman and E. J. Johnson, *The Adaptive Decision Maker*, Cambridge University Press, Cambridge, UK, 1993.
28. H. Montgomery and O. Svenson, A Think Aloud Study of Dominance Structuring in Decision Processes, *Process and Structure in Human Decision Making*, 1989, pp. 135–150.
29. R. S. Billings and S. A. Marcus, Measures of Compensatory and Noncompensatory Models of Decision Behavior: Process Tracing Versus Policy Capturing, *Organizational Behavior and Human Performance*, **31**(3), 1983, pp. 331–352.
30. A. Glockner and T. Betsch, Multiple-reason Decision Making Based on Automatic Processing, *Journal of Experimental Psychology: Learning, memory, and cognition*, **34**(5), 2008, pp. 1055–1075.
31. A. Glockner and A. Herbold, An Eye-Tracking Study on Information Processing in Risky Decisions: Evidence for Compensatory Strategies based on Automatic Processes, *Journal of Behavioral Decision Making*, **24**(1), 2011, pp. 71–98.
32. J. H. Goldberg and X. P. Kotval, Computer Interface Evaluation using Eye Movements: Methods and Constructs, *International Journal of Industrial Ergonomics*, **24**(6), 1999, pp. 631–645.
33. G. R. Loftus and N. H. Mackworth, Cognitive Determinants of Fixation Location During Picture Viewing, *Journal of Experimental Psychology: Human Perception and Performance*, **4**(4), 1978, pp. 565–572.
34. A. Madsen, Studies of Visual Attention in Physics Problem Solving, Ph.D. thesis, Kansas State University, Manhattan, KS, USA, 2013.
35. M. I. Posner, C. R. Snyder and B. J. Davidson, Attention and the Detection of Signals, *Journal of Experimental Psychology*, General **109**(2), 1980, pp. 160–174.
36. K. Rayner, Eye Movements in Reading and Information Processing: 20 Years of Research, *Psychological Bulletin*, **124**(3), 1998, pp. 372–422.
37. J. W. Schoonard and J. D. Gould, Field of View and Target Uncertainty in Visual Search and Inspection, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, **15**(1), 1973, pp. 33–42.
38. A. D. Smith, J. P. Mestre and B. H. Ross, Eye-gaze Patterns as Students Study Worked-Out Examples in Mechanics, *Physical Review Special Topics-Physics Education Research*, **6**(2), 2010, pp. 020118-1-020118-9.
39. M. Hegarty, R. E. Mayer and C. E. Green, Comprehension of Arithmetic Word Problems: Evidence from Students' Eye Fixations, *Journal of Educational Psychology*, **84**(1), 1992, pp. 76–84.
40. L. Verschaffel, E. De Corte and A. Pauwels, Solving Compare Word Problems: as Eye Movement Test of Lewis and Mayer's Consistency Hypothesis, *Journal of Educational Psychology*, **84**(1), 1992, pp. 85–94.
41. A. Gegenfurtner, E. Lehtinen and R. Slj, Expertise Differences in the Comprehension of Visualizations: A Meta-Analysis of Eye-Tracking Research in Professional Domains. *Educational Psychology Review*, **23**(4), 2011, pp. 523–552.
42. A. M. Madsen, A. M. Larson, L. C. Loschky and N. S. Rebello, Differences in Visual Attention between Those Who Correctly and Incorrectly Answer Physics Problems, *Physical Review Special Topics-Physics Education Research*, **8**(1), 2012, pp. 010122-1–010122-13.
43. M.-J. Tsai, H.-T. Hou, M.-L. Lai, W.-Y. Liu and F.-Y. Yang, Visual Attention for Solving Multiple-Choice Science Problem: An Eye-Tracking Analysis, *Computer and Education*, **58**(1), 2012, pp. 375–385.
44. D. Denick, A. I. Santiago-Romn, R. A. Streveler and N. Barrett, Validating the Diagnostic Capabilities of Concept Inventories: Preliminary Evidence from the Concept Assessment Tool for Statics (CATS), *In Proceedings of the 2012 American Society of Engineering Education Conference and Exposition*, San Antonio, TX, 10–13 June, 2012.
45. P. S. Steif and J. A. Dantzler, A Statics Concept Inventory:

Development and Psychometric Analysis, *Journal of Engineering Education*, **94**(4), 2005, pp. 363–371.
46. P. S. Steif and M. Hansen, Comparisons between Performances in A Statics Concept Inventory and Course Examinations, *International Journal of Engineering Education*, **22**(5), 2006, pp. 1070–1076.
47. P. S. Steif, Initial Data from A Statics Concept Inventory, *In Proceedings of the 2004 American Society of Engineering Education Conference and Exposition*, Salt Lake City, UT, 20–23 June 2004.
48. R. Guay, Purdue Spatial Visualization Test. Purdue University, 1976.
49. S. Y. Yoon, Psychometric Properties of the Revised Purdue Spatial Visualization Tests: Visualization of Rotations (The Revised PSVT-R), Ph.D. Diss., Purdue University, 2011.
50. T. J. Branoff, Spatial Visualization Measurement: A Modification of the Purdue Spatial Visualization Test-Visualization of Rotations, *Engineering Design Graphics Journal*, **64**(2), 2000, pp. 14–22.
51. Index of Learning Styles Questionnaire, http://www.engr.ncsu.edu/learningstyles/ilsweb.html, Accessed on December 2015.
52. R. M. Felder and L. K. Silverman, Learning and Teaching Styles in Engineering Education, *Engineering Education*, **78**(7), 1988, pp. 674–681.
53. R. M. Felder and J. Spurlin, Applications, Reliability and Validity of the Index of Learning Styles, *International Journal of Engineering Education*, **21**(1), 2005, pp. 103–112.
54. K. Pernice and J. Nielsen, Eyetracking Methodology: How to Conduct and Evaluate Usability Studies Using Eyetracking, *Nielsen Norman Group Technical Report*, 2009.
55. The polychromatic version of Fig. 2, https://engineering.purdue.edu/reidlab/pdf/IJEE2016Fig2.pdf, Accessed on October 2016.
56. B. D. Coller, An Experiment in Hands-On Learning in Engineering Mechanics: Statics, *International Journal of Engineering Education*, **24**(3), 2008, pp. 545–557.
57. K. M. Steele, S. R. Brunhaver and S. D. Sheppard, Feedback from in-class Worksheets and Discussion Improves Performance on the Statics Concept Inventory, *International Journal of Engineering Education*, **30**(4), 2014, pp. 992–999.
58. A. Dollar and P. S. Steif, Learning Modules for Statics, *International Journal of Engineering Education*, **22**(2), 2006, pp. 381–392.
59. M. Borrego, J. E. Froyd, C. Henderson, S. Cutler and M. Prince, Influence of Engineering Instructors' Teaching and Learning Beliefs on Pedagogies in Engineering Science Courses, *International Journal of Engineering Education*, **29**(6), 2013, pp. 1456–1471.
60. M. L. Lai, M. J. Tsai, F. Y. Yang, C. Y. Hsu, T. C. Liu, S. W. Y. Lee and C. C. Tsai, A Review of Using Eye-Tracking Technology in Exploring Learning from 2000 to 2012, *Educational Research Review*, **10**, 2013, pp. 90–115.
61. N. Fang and J. Lu, A Decision Tree Approach to Predictive Modeling of Student Performance in Engineering Dynamics, *International Journal of Engineering Education*, **26**(1), 2010, pp. 87–95.
62. S. C. Chen, H. C. She, M. H. Chuang, J. Y. Wu, J. L. Tsai and T. P. Jung, Eye Movements Predict Students' Computer-Based Assessment Performance of Physics Concepts in Different Presentation Modalities, *Computers & Education*, **74**, 2014, pp. 61–72.
63. D. Li, J. Babcock and D. J. Parkhurst, openEyes: A Low-Cost Head-Mounted Eye-Tracking Solution, *In Proceedings of the 2006 ACM symposium on Eye tracking research & applications*, 2006, pp. 95–100.
64. J. San Agustin, H. Skovsgaard, E. Mollenbach, M. Barret, M. Tall, D. W. Hansen and J. P. Hansen, Evaluation of A Low-Cost Open-Source Gaze Tracker, *In Proceedings of the 2010 ACM Symposium on Eye-Tracking Research & Applications*, 2010, pp. 77–80.
65. M. Kassner, W. Patera, and A. Bulling, Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-Based Interaction, *In Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication*, 2014, pp. 1151–1160.
66. A. D. R. I. A. N. Voßkühler, V. Nordmeier, L. Kuchinke and A. M. Jacobs, OGAMA (Open Gaze and Mouse Analyzer): Open-Source Software Designed to Analyze Eye and Mouse Movements in Slideshow Study Designs, *Behavior Research Methods*, **40**(4), 2008, pp. 1150–1162.
67. D. Litchfield, L. Ball, T. Donovan, D. J. Manning and T. Crawford, Viewing Another Person's Eye Movements Improves Identification of Pulmonary Nodules in Chest X-Ray Inspection, *Journal of Experimental Psychology*, **16**, 2010, pp. 251–262.
68. T. van Gog, H. Jarodzka, K. Scheiter, P. Gerjets and F. Paas, Attention Guidance during Example Study via the Model's Eye Movements, *Computers in Human Behavior*, **25**, 2009, pp. 785–791.

Youyi Bi is a Ph.D. candidate in the School of Mechanical Engineering, Purdue University. He obtained his B.S. in Mechanical Engineering and M.S. in Aerospace Manufacturing Engineering both from Beihang University, China. His research interests include engineering education, decision-making and optimization in mechanical design, psychoacoustics in product design and design theory and methodology.

Tahira Reid is an Assistant Professor in the School of Mechanical Engineering at Purdue University and is the director of the Research in Engineering and Interdisciplinary Design (REID) Laboratory. Her research interests include developing methods that help engineers think critically about perception-based factors and their impact on engineered design solutions. Prior to arriving at Purdue in 2012, she completed a postdoctoral position in the Mechanical Engineering department at Iowa State working in the Interdisciplinary Research in Sustainable (IRIS) Design Laboratory under the mentorship of Erin MacDonald. In 2010, she received her PhD from the University of Michigan in Design Science, with Mechanical Engineering and Psychology as her focus areas. Dr. Reid received both her BS and MS degrees in Mechanical Engineering from Rensselaer Polytechnic Institute in 2000 and 2004, respectively. Some of her early accomplishments and research have been featured on NBC's Today Show, National Public Radio (NPR), the Washington Post, Smithsonian Magazine, and many other media outlets.