USING WORD CLOUDS FOR FAST, FORMATIVE ASSESSMENT OF STUDENTS’ SHORT WRITTEN RESPONSES

Bill J. Brooks*, Debra M. Gilbuena*, Stephen J. Krause**, and Milo D. Koretsky*

* Oregon State University • Corvallis, OR 97331-2702
** Arizona State University • Tempe, AZ 85281

Bill Brooks is a postdoctoral scholar in the School of Chemical, Biological, and Environmental Engineering at Oregon State University. He received his B.S. and his Ph.D. from Oregon State University, both in chemical engineering. His Ph.D. research involved the use of written explanations to concept questions to investigate technology-mediated active learning in the undergraduate chemical engineering classroom. He is interested in using technology to enhance educational practices in promoting conceptual understanding. He is the primary programmer of the AIChE Concept Warehouse and his current focus is on its continued development, specifically creating and integrating Interactive Virtual Labs.

Debra Gilbuena is a postdoctoral scholar in the School of Chemical, Biological, and Environmental Engineering at Oregon State University. She received her Ph.D. from Oregon State University in chemical engineering with a dissertation focused on engineering education. Debra also has received B.S., M.S., and MBA degrees from OSU. She has 4 years of industrial experience including a position in sensor development, an area in which she holds a patent. Her research currently has two focus areas: 1) the characterization and analysis of feedback, student learning, and engagement in project-based learning, and 2) the diffusion of effective educational interventions, materials, and practices.

Stephen Krause is a professor in the Materials Science Program in the Fulton School of Engineering at Arizona State University. He teaches in the areas of introductory materials engineering, polymers and composites, and capstone design. His research interests include evaluating conceptual knowledge, misconceptions, and technologies to promote conceptual change. He has co-developed a Materials Concept Inventory and a Chemistry Concept Inventory for assessing conceptual knowledge and change for introductory materials science and chemistry classes. He is currently conducting research on strategies that use Internet tools and resources that promote conceptual change for enhancement of students’ attitudes, achievement, and persistence. He was a co-author for best paper awards from the Frontiers in Education Conference in 2009 and the Journal of Engineering Education in 2013.

Milo Koretsky is a professor in the School of Chemical, Biological, and Environmental Engineering at Oregon State University. He received his B.S. and M.S. degrees from UC San Diego and his Ph.D. from UC Berkeley, all in chemical engineering. He is interested in integrating technology into effective educational practices and in promoting the use of higher-level cognitive skills in engineering problem solving. His research interests particularly focus on what prevents students from being able to integrate and extend the knowledge developed in specific courses in the core curriculum to the more complex, authentic problems and projects they face as professionals.

Imagine your department chair has just assigned you to teach Material and Energy Balances, a required course that has grown considerably in enrollment for the last several years. You taught it a few years ago and used ConcepTests® for in-class active learning with reasonable success. You plan to use them again this time around. You recently attended a professional development seminar that described the learning benefits of asking your students to write explanations and reflections. It sounds like a great idea, so you decide to have your students write explanations to justify their answer choices to their ConcepTests. You try it the first week of class. After class you are checking out the responses, plotting the answer distributions, and then it hits you. You see the 250 written explanations. It is going to take hours to read and analyze all of these explanations! If you don’t read them, will your students take them seriously? Will they continue to reflect and get the most out of them? If you do take the time to read all of them, what are you sacrificing? What part of your class preparation are you giving up? What do you do?

Many instructors have approached a dilemma similar to the one discussed in the vignette above. Sometimes it happens when you are first contemplating the implementation of a research-based instructional strategy and sometimes it comes, as in the vignette, only after the first implementation. This article presents a potential solution to the vignette dilemma of analyzing short written responses—the use of word clouds. Word clouds provide a visual representation of word usage and frequency. They offer a quick visualization of aggregate text responses to reduce the burden of information overload. When combined with an audience response system, they afford instructors a way to easily analyze written explanations from tens or hundreds of students in a very short time.

At right is a word cloud summary of this article.

Can you figure out the main point from the prominent words in the word cloud?
In this article, we describe how word clouds can be used for formative assessment in active learning. In particular, we discuss how they have been integrated into and used with the AIChE Concept Warehouse (CW). The CW is a web-based tool to help the chemical engineering education community more easily use active learning pedagogies.[3] We focus on the ways word clouds afford improved instruction through the CW through their use as a formative assessment tool that can provide instructors and students with valuable, timely feedback. We illustrate the use of word clouds with evidence from two active-learning examples: student in-class responses to multiple-choice concept questions, during the first part of peer instruction[13]; and student responses to “muddiest point” reflection exercises. Word clouds have been used as a qualitative analysis tool in teaching. Ramsden and Bate[25] suggest the use of word clouds as a complementary method to other research and teaching methods.

**BACKGROUND**

Active-learning pedagogies have been shown to improve student conceptual understanding.[6] Active learning means more than engaging students in classroom exercises; activities should be designed around learning outcomes, promote student reflection, and get students to think about what they are learning.[7] Formative assessment is one integral aspect of these pedagogies that helps meet these design criteria. Assessments that include students’ short written explanations[4,5,9-11] can enhance learning.[8,12-15] However, it is difficult to expediently examine written responses in large classes.

Word clouds, also known as “tag clouds” or “term clouds,” can be a useful analytical tool to summarize text data and provide meaningful interpretations.[10] Word clouds have been found to be beneficial because they are “highly interpretable,” giving a direct visual representation of the content being measured.[17] They have been used both as a research tool and as a teaching tool.

As a research tool, McNaught & Lam[18] showed how word clouds can uncover themes in interviews consistent with those identified by other qualitative analysis methods. Similarly, word clouds have been used as a qualitative analysis tool in other cases.[19,20] While word clouds are generally interpreted in terms of the most common words, attention to missing words or infrequent words can be just as important.[17,21] The context from which a word cloud is created also plays an important role in the interpretation of the resultant word clouds.[17] e.g., the phrase “energy balance” holds one meaning in a chemical engineering course and takes on an entirely different meaning in Oriental medicine.[22]

Educators have also begun to report the benefits of word clouds in teaching. Ramsden & Bate[23] put forth a general working paper presenting word clouds as a useful teaching tool comparing different word cloud software and discussing aspects educators should consider. They note that data needs to be in a usable state (i.e., as electronic text) for word cloud analysis. In addition, they note the following potential limitations of word cloud software: spelling errors may not be taken into account; words that appear to be common may be eliminated even if they represent an important acronym, e.g., IT; word clouds represent frequency, not necessarily importance; and word clouds often fail to group similar words. Ramsden and Bate[25] suggest the use of word clouds as a complementary method to other research and teaching methods.

In practice, educators have described having students construct word clouds from pre-existing materials (such as speeches, quotes, and web pages) to summarize and promote reflection and discussion in many fields, including: accounting,[24] social studies,[25] teaching vocabulary,[26] and theology.[27] Word clouds have also been used in several ways for teaching reading and writing.[28] In one example that resembles our use of “muddiest point” reflections, an instructor used word clouds to summarize students' text messages in a high school English class to formatively assess what was learned in the previous class.[29] In this article, we present how word clouds can be used in chemical engineering education. We illustrate how the CW affords automatic aggregation of students’ writing and word cloud construction. This system eliminates the need to manually collect and transcribe handwritten reflections in order to construct word clouds.

**AIChE CONCEPT WAREHOUSE**

The CW was used as the primary data collection tool for the examples reported in this article. It is a database-driven website facilitating the use of concept questions throughout the core chemical engineering curriculum. Currently the CW has more than 2,000 concept questions (ConcepTests) and 10 valid and reliable concept inventories available for searching, viewing, and using in courses. Instructor and student interfaces are available for use at <http://cw.edudiv.org>, and university faculty can obtain an account through this site. More general information about this tool can be found elsewhere.[28] In this article, we focus on the word cloud feature that facilitated formative assessment.

For context, we provide a detailed description of the algorithm used to generate word clouds in the CW. A wide variety of word cloud algorithms are reported in the literature, e.g., some count the frequency of individual words while others count frequency of word pairs.[30] Currently the CW summarizes the frequency of single words only. To generate the word clouds, the CW first aggregates all of the written explanations into a single string of text per answer choice. It removes HTML tags like ‘br’ and ‘quot’ as well as filler words (e.g., able, about, above, according, accordingly, across, actually,…). The words as they were submitted temporally are mapped into the cloud horizontally and vertically according to English convention. Word frequency is mapped to the
word size and color. Blue, bigger words are more frequent. Red, smaller words are less frequent. While some criticize word clouds for ignoring semantic relationships such as similar words,[17,23] the algorithm the CW uses has been improved to ignore case differences and combines similar words like singular and plural forms.

**EXAMPLE 1: WORD CLOUDS IN CONCEPTTEST ASSESSMENT**

The first active-learning example we use illustrates the use of word clouds in the context of polling during peer instruction. Peer instruction is arguably the most well-known and widely used technology-mediated active-learning pedagogy in post-secondary Science, Technology, Engineering, and Mathematics education.[1,31] It consists of a structured polling process where a concept question (also called a ConcepTest or ‘clicker’ question) is presented to the class. Students first answer the question individually. They are then encouraged to discuss the answer choices in small groups. Finally, they individually submit a final answer. This sequence is then typically followed by a class-wide discussion. The data presented below comes from the first individual answering where students are asked to explain in writing their choice for two sample concept questions.

**Methods**

Data were collected from two cohorts enrolled in a required, sophomore-level, undergraduate energy balances course at a large public university. Between 60 to 70 chemical, biological, and environmental engineering students provided written explanations for each of the two questions presented in this article. These students came from a subset of a larger study population, reported elsewhere.[8] The lectures and recitations for both cohorts were taught by the same instructor, in the same room, using the CW to deliver the ConcepTests. The Institutional Review Board approved the research and participants signed informed consent forms.

Figure 1 and Figure 2 depict two isomorphic concept questions as they were presented to the students in their respective cohort; one cohort answered one question, the other cohort answered the other question. In isomorphic questions, students need to apply the same core

**Figure 1. Sample question (piston question) as it was presented to students who wrote explanations.**

**Figure 2. Sample question (balloon question) as it was presented to students who wrote explanations.**

Air at high pressure and ambient temperature is contained in a perfectly insulated piston-cylinder device. If the locks holding the piston in place are removed, the piston moves upwards to a stopper. The temperature of the air _____________.

<table>
<thead>
<tr>
<th>increases</th>
<th>remains the same</th>
<th>decreases</th>
</tr>
</thead>
</table>

Please explain your answer in the box below.

Please rate how confident you are with your answer.

- substantially unsure
- moderately unsure
- neutral
- moderately confident
- substantially confident

Submit

A perfectly insulated balloon filled with an ideal gas rises into the sky. As the balloon rises, the external pressure decreases, causing the balloon to expand. What happens to the temperature of the gas inside the balloon?

<table>
<thead>
<tr>
<th>increases</th>
<th>remains the same</th>
<th>need more information</th>
<th>decreases</th>
</tr>
</thead>
</table>

Please explain your answer in the box below.

Submit
concept, but the questions have different surface features, like the balloon or piston in these questions, or what Smith, et al.[32] calls “different cover stories.” To answer the questions correctly, ideally, students apply their knowledge of energy balances, recognizing that the work done by the gas on the surroundings lowers its internal energy and, therefore, its temperature. For this question pair, however, the correct answer can also be obtained from faulty reasoning using the ideal gas law. In that case a student may reason, since $PV = nRT$, as $P$ decreases, $T$ also must decrease. This reasoning process fails to account for changing volume, and is, therefore, classified as faulty reasoning.

Results

Figures 3 and 4 present the word clouds for the questions depicted in Figures 1 and 2, respectively. The left-hand side of each figure shows the answer options for the respective question. The middle contains the word cloud that was produced from the aggregation of written explanations for the corresponding answer option. The right-hand side contains a representative explanation given by a student that selected the corresponding answer. All written explanations for both questions were also iteratively coded, detailed results of which are presented elsewhere.[8]

Students predominantly chose the correct answer for the piston question (Figure 1) using scientifically valid reasoning, yet students who answered the balloon question (Figure 2) correctly predominantly used faulty reasoning related to the ideal gas law. For the balloon question most students chose “remains the same.” They did so apparently thought that since the balloon was “perfectly insulated”—no heat meant there could be no temperature change. Students who select this answer to the two conceptually similar questions do so using different reasons, and the word clouds capture this difference.

So what can we learn from these word clouds?

First, let us focus on the correct answer, “decreases.” The students who predominantly chose the correct answer for the piston question using correct reasoning have a corresponding word cloud in which the words “energy” and “work” can be seen. However, students who answered the balloon question correctly using faulty reasoning related to the ideal gas law, have a corresponding word cloud in which “energy” is present but “work” does not appear; terms like “pv” and “nrt” can be seen instead. In the case of the balloon problem, we see an example of when a missing word is as important as, or more so than, the words that appear.[17,21] In this case, the word cloud without the word “work” suggests that even though many students chose correctly, they may still need attention regarding the role of work in closed-system energy balances.

We can also consider the word clouds associated with explanations for the distractors to provide insight into the ideas expressed by students who chose a wrong answer. The students that answered the piston question with “remains the same” thought that the decrease in pressure was compensated apparently thought that the decrease in pressure was compensated for by the increase in volume. Students who select this answer to the two conceptually similar questions do so using different reasons, and the word clouds capture this difference.

![Figure 3. Multiple-choice answer options, word clouds, and representative explanations for the concept question depicted in Figure 1 (piston question).](image)
for by the increase in volume. Notice that the words “pressure,” “volume,” “decrease,” and “increase” are in almost equal proportion. For the balloon question, most students chose “remains the same” since the balloon was “perfectly insulated”—no heat meant there could be no temperature change. Therefore, neither “energy” nor “work” appear in that word cloud. Again, the reasons for similar answers are different. In the case of the distractors, each distractor should have a particular misconception with which it is most associated, but it could have several. When combined with the instructor’s expertise and previous experience, a word cloud may provide an instructor with enough information to identify which misconception is most prevalent for the majority of students who chose each distractor.

To further help identify misconceptions, an instructor can click on any particular word in the word cloud to easily view the subset of the explanations that contain that word. For example, Figure 5 (page 196) shows the screenshot of the page that results when the hyperlink for the word “weight” was selected from the word cloud associated with the “decrease” answer option in Figure 3. Both a filtered word cloud and a list of all explanations where students used the word “weight” are shown so that instructors can focus on a set of explanations to consider. In this case, it is evident that the five students using the word “weight” correctly and explicitly associated the expansion work of the gas with the movement of the weight. Most of those students are also explicitly connecting the work done by the air on the weight to the decrease in internal energy, demonstrating that they properly applied an energy balance. This feature allows instructors to ascertain how students used a particular word and what concepts associated with the word are elicited in their reasoning.

### EXAMPLE 2: WORD CLOUDS TO EXAMINE MUDDIEST POINT REFLECTION

The second example illustrates the use of Muddiest Point Reflections for formative assessment. In a Muddiest Point Reflection, an instructor asks students to write a brief, anonymous written comment describing the concept or topic that they found to be the most difficult to understand during class. With this information, the instructor can strategize to adjust his/her teaching and pedagogy to address issues specific to many students. The CW software allows word clouds of Muddiest Point Reflections to be available shortly after students have responded. It also provides links to words that allow filtered word clouds analogous to that shown in Figure 5.

<table>
<thead>
<tr>
<th>Answer Option</th>
<th>Word Cloud of Written Explanations</th>
<th>Representative Explanation (emphasis added)</th>
</tr>
</thead>
<tbody>
<tr>
<td>remains the same</td>
<td>balloon perfectly insulated remain heat increase volume temperature decrease change gas pressure pv nrt inside</td>
<td>the balloon is perfectly insulated so the temperature of the balloon does not change.</td>
</tr>
<tr>
<td>increases</td>
<td>pv=nrt, v is increase, so t is increase</td>
<td></td>
</tr>
<tr>
<td>decreases</td>
<td>Temperature must go down to maintain PV=nRT relationship.</td>
<td></td>
</tr>
<tr>
<td>need more information</td>
<td>PV=nRT, or T=PV/(nR). Because P is decreasing, but V is increasing, we need to know how exactly they are related in order to know if temperature is increasing or decreasing.</td>
<td></td>
</tr>
</tbody>
</table>
Methods

Data were collected in several materials science classes at a large public university with class sizes of 40–45 students. Figure 6 presents a screenshot of the Muddiest Point Reflection as it is presented to students on their laptops, cell phones, or tablets using the CW. The Muddiest Point Reflection was assigned at the end of class and students could answer on their electronic devices; however, the assignment was allowed to be submitted up to six hours after class. Students were offered up to five percent extra credit on their final grade for answering at least 20 of the 24 Muddiest Point Reflections over the semester. These exercises have an estimated 65% response rate. In addition to the Muddiest Point Reflections, 33 students from one section answered a survey about the impact of word cloud use in the classroom. The data collected for this research was approved by the Institutional Review Board.

When the exercise was first presented to students, the instructor discussed with students the purpose of the exercise, both from a student learning and an instructor feedback standpoint. At the beginning of the class following each Muddiest Point Reflection submission, the instructor thanked the students for their submissions. In addition, the instructor showed the single word cloud aggregated from all student responses to the previous submission, presented student quotes, and led a discussion regarding the student learning issues. The discussion used the method of Socratic questioning in working toward resolution of the student learning issues. The instructor also reiterated that responses to the Muddiest Point Reflection would help improve not only the course for the current cohort of students, but for future cohorts as well.

Results

Figure 7 presents the resultant word cloud from an aggregate of all students’ Muddiest Point Reflections after the topic of failure in metals was covered in class. Figure 7 also includes representative quotes. This topic has important real-world consequences, since engineering systems such as airplanes, chemical plants, and bridges are susceptible to failures with consequent loss of lives.

So what can we and students learn from the Muddiest Point Reflection word cloud?

In the prior class discussion of this topic, the four main types of failures were described, along with the failure mechanisms, fracture appearances, and testing methods that have predictive capabilities. “Failure,” “mechanism,” and “types” were the largest words seen, indicating that failure types and associated mechanisms were the most prominent muddiest points as opposed to fracture appearances or testing methods. The major difficulty that a significant fraction of the students were grappling with was the connections between the different aspects of a given “type of failure mechanism,” which was clearly reflected in the size of the words in the word cloud.

A reading of the student comments confirmed the diagnosis that was first quickly highlighted by the word cloud. Because of this information, the instructor was inspired in the next class to create a well-detailed table delineating the characteristics of the failure mechanism types and features. The table included: a real-world example, conditions causing failure, mechanism of failure, fracture surface appearance, and test methods for predicting lifetime associated with different mechanisms. Most of the students vigorously took notes and copied the table during the discussion. This example illustrates how the use of word clouds in Muddiest Point Reflections helps the instructor improve and adjust instruction. The rapid feedback with the Muddiest Point Reflections and associated word cloud can have a significant impact on student learning.

Research has shown that addressing learning issues as quickly as possible with rapid feedback is very effective for improving motivation and learning.[34]

![Figure 5. Sample filtered word cloud from the explanations aggregated into and summarized by the word clouds in Figure 3. They are limited to only the explanations that used the word “weight.”](image-url)
Frequent feedback plays an important role in the progression of a learner from the level of “novice” toward “expert” understanding and performance in a given domain. In a review on the acquisition of expert skills, Ericsson, et al.[35] cite one important condition for optimal learning and improving performance is that learners will receive immediate and informative feedback and knowledge of results of their performance on a given task. This is reflected by the response of the students to a survey about the use of word clouds.

Thirty-three students participated in an end-of-semester survey about the impact of the Muddiest Point Reflection word clouds in the classroom. Sixty-seven percent of those students agreed or strongly agreed that, “The word clouds helped me visualize what the most confusing concepts in the class were.” Seventy-six percent of the students agreed or strongly agreed with the statement that, “The word clouds informed me about issues other students were having with the class content.” For instructors the word clouds and Muddiest Point Reflection provide a quick and measured diagnosis of student learning issues for adjusting current and future instruction. For students, the word clouds serve as a visual indicator of issues that they and others in class may be grappling with and they are more motivated to engage in discussion and dialogue in addressing those issues to improve their knowledge and learning on more difficult concepts and content. Thus, instructors and students are mutual beneficiaries of the use of word clouds in materials classes.

**Figure 6.** Muddiest point reflection as it was delivered to students.

**Figure 7.** Word cloud and sample quotes for the sample Muddiest Point Reflection.

**WORD CLOUDS FOR OTHER SHORT WRITTEN EXERCISES**

In the previous two sections we discussed how word clouds have been used for specific types of exercises, ConcepTests and Muddiest Point Reflections. While further research is required to evaluate the utility of word clouds to examine other types of short written exercises, in this section we briefly explore a few other areas where word clouds may be beneficial. In general, word clouds can be used for any type of short written response. For example, Vigeant, Prince, and Nottis[36] describe inquiry-based activities for thermodynamics and heat transfer. In these activities students are prompted to predict results of an experiment before the experiment and explain their prediction in writing. The authors then have students run or observe an experiment. After experimentation, students compare results with their predictions in writing, and discuss with their peers. Finally they write answers to post-activity questions. Each of the writing steps presents an opportunity for word cloud use to visualize aggregate student responses. As their inquiry-based activities continue to be implemented at different schools in different contexts, word clouds might offer another quick way to examine if the students in these new contexts give similar responses to those in the original context. Other scaffolded activities that have a similar “predict - observe - explain” structure[37] such as the interactive virtual laboratories recently incorporated in the CW, may also benefit from word clouds.

**PLANNED WORD CLOUD IMPROVEMENTS FOR THE AICHE CONCEPT WAREHOUSE**

For exercises like Muddiest Point Reflections and other short written exercises, the current word cloud analytical algorithm may be sufficient. However, this type of analysis may benefit from including the option of using word pairs[30] as a basis, an option we are currently exploring. Word pairs, if they maintain word order, might highlight instances where the order of the words is as critical as the individual words themselves. In addition, we are considering modifications to address one of the concerns reported by Ramsden and Bate[23]: to prevent the elimination of seemingly common words with special meanings (it versus IT) we intend to incorporate a custom list in which instructors can identify words to exclude or words to include.

**CONCLUSIONS AND IMPLICATIONS**

Active learning can help students develop deeper understanding of chemical engineering principles. While multiple-choice ConcepTests...
are useful, we advocate for including student writing in learning activities as well. Writing explanations and reflections can help students organize their thinking and explicitly reflect on what has been covered. These types of assignments also provide information that faculty can use to focus instruction. Examining students’ writing reveals their faulty reasoning and misconceptions, and can help the instructor identify concepts and topics that are difficult.

In this article, we demonstrate that word clouds can provide a quick analytical technique to begin assessment of student written explanations and reflections. The AIChe Concept Warehouse automatically generates word clouds and can facilitate collection and analysis of student writing, even in large classes. Unlike external applications like Wordle where text needs to be manually entered, word clouds are automatic and quick. However, even with automation, instructors still face the challenge of interpreting the information provided; consequently, we are also working to improve and better integrate word clouds and include other analysis options into the AIChe Concept Warehouse. Our goal is to help make deep, concept-based learning more effective for students and easy for faculty to implement.

ACKNOWLEDGMENTS
The authors would like to acknowledge funding from the National Science Foundation (DUE 1023099, 1225456, 1226325) and from Oregon State University (TRF 931, 1056). The opinions expressed are strictly those of the authors and do not necessarily represent those of the National Science Foundation.

REFERENCES
23. Ramsden, A., and A. Bate, “Using Word Clouds in Teaching and Learning,” University of Bath (Unpublished) 2008 <http://opus.bath.ac.uk/474/1/using%252520word%252520clouds%252520in%252520teach%2520on%2520in%2520learning.pdf> Retrieved 03/11/14