Comprehension SEEDING: Comprehension through *S*elf *E*xplanation, *E*nhanced Discussion, and *IN*quiry *G*eneration

Frank Paiva[1], James Glenn[1], Karen Mazidi[1], Robert Talbot[2], Ruth Wylie[3], Michelene T.H. Chi[3], Erik Dutilly[4], Brandon Helding[5], Mingyu Lin[1], Susan Trickett[5], and Rodney D. Nielsen[1]

[1] University of North Texas
{frankpaiva,jamesglenn2,karenmazidi,mingyulin}@my.unt.edu,
rodney.nielsen@unt.edu
[2] University of Colorado Denver
robert.talbot@ucdenver.edu
[3] Arizona State University
{ruth.wylie,michelene.chi}@asu.edu
[4] University of Colorado Boulder
erik.dutilly@colorado.edu
[5] Boulder Language Technologies
b.a.helding@gmail.com, sbtrickett@bltek.com

Abstract. In this paper we introduce the Comprehension SEEDING system and describe the system components designed to enhance classroom discussion by providing real-time formative feedback to teachers. Using SEEDING, teachers ask free-response questions. As students are constructing their responses using digital devices, SEEDING allows teachers to assess a student's understanding. Once SEEDING collects student responses, the system automatically groups them based on semantic similarity. Teachers can use this information to address student misconceptions and engage the classroom from a more informed perspective. This paper describes the SEEDING system and how it can be used to aid teachers and improve classroom discussion.

1 Introduction

Teachers ask students questions in the classroom both to assess their understanding and also to facilitate learning. Students learn as a result of engaging with the material and participating in shared discourse (Larson, 2000). Although this can potentially be a reasonable way to generate classroom discussion, effective classroom engagement is difficult to achieve this way because teachers can only involve one student at a time. This may cause other students to become disengaged from the discussion. To address this problem, classroom response technologies such as clickers, have been shown to improve student learning and engagement by allowing all students to answer, while providing the teacher with real-time formative feedback.

S. Trausan-Matu et al. (Eds.): ITS 2014, LNCS 8474, pp. 283–293, 2014.

Clickers are a classroom response system in which each student has a handheld remote control by which they respond to questions that are projected onto a screen in the classroom. Previous work on clickers has shown that they can be beneficial for enhancing student learning and engagement (Duncan, 2006; Fies & Marshall, 2006; Herreid, 2006; Keller et al., 2007; Penuel, Boscardin, Masyn, & Crawford, 2006; Siau, Nah, Siau, Sheng, & Nah, 2006). However, there are limitations that could explain why small-scale efficacy tests for the use of the technology have seen mixed results (Bunce et al., 2006; Carnaghan & Webb, 2007; Duggan et al., 2007). In order for teachers to take advantage of clickers and any automated response tallying, teachers are limited to asking multiple choice questions. Although multiple choice questions are helpful when assessing basic factual knowledge, it can be difficult to assess deep knowledge in a closed-response question format (Campbell, 1999; McNeill et al. 2009). The effectiveness of clickers is limited to the quality of the multiple-choice questions that teachers pose, and it is difficult and time consuming to construct questions with good distractors. Even with meaningful distractors, multiple-choice questions only require students to *recognize*, rather than *generate* the correct response. According to the Interactive, Constructive, Active, Passive (ICAP) framework (Chi, 2009), constructing answers to free-response questions is a more cognitively engaging task than simply selecting answers to multiple-choice questions and should result in deeper learning.

One of SEEDING's goals is to improve on the engagement advantages afforded by clickers, while largely eliminating their weaknesses. Specifically, SEEDING is a new classroom learning technology that: allows teachers to pose free-response questions, results in all students constructing responses, provides teachers real-time formative feedback, and aims to encourage deeper questions in the classroom.

2 Comprehension SEEDING

SEEDING is grounded in results from three key areas of cognitive and learning sciences research, 1) student self explanation, 2) formative assessment with classroom engagement and discourse, and 3) educational question asking practices. The Comprehension SEEDING system is divided into three analogous distinct but related components that work together to create an enhanced learning environment for both teachers and students. These three components, self-explanation (SE), enhanced discussion (ED), and inquiry generation (ING), are summarized in this section and detailed in the sections that follow, while highlighting their theoretical advantages.

The Comprehension SEEDING system allows teachers to pose free-response questions. Students answer these questions via digital devices (each of the students in our current study, approximately 1250 in total, is using a Google Nexus 7, but classrooms outside the study have used laptops, netbooks, various tablets, android phones, iPhones, and other digital devices). While students compose their responses, the system provides a real-time analysis of the student responses.

Once SEEDING receives most of the student responses, it automatically groups them in up to four clusters. Teachers have the option to view and share each student response with the class. However, showing individual responses can be time consuming and may address misconceptions only held by a few students. Using the clusters, teachers can quickly determine the current overall status of the classroom's understanding of the question posed.

SEEDING allows each student and teacher to interact with the current presented material. To achieve this level of individual interaction, the system needs to address the different requirements of the teacher versus the students and allow each student to use the system simultaneously in the classroom. Our approach consists of a web-based solution that in the present study, runs on Nexus 7 tablets for the students and typically runs on a desktop or laptop for the teacher.

SEEDING operates differently based on the user's role (e.g, student, teacher). Teachers using the system use their classroom computer which connects to a projector. This provides the teacher with two windows, a control dashboard and a classroom display. The first control window, gives teachers the ability to control, manage, view, assess, and teach the classroom. The second window allows the teacher to share student responses, vocabulary words, and images with the classroom. Unlike the teacher windows, we have provided a minimal interface for the students – they can log in, receive questions & vocabulary words, construct their responses, and logout. As an alternative to using the on-screen keyboard, students' tablets are complimented with a physical keyboard to reduce student response time during classroom sessions

3 Self Explanation

Self Explanation. Given a question, Comprehension SEEDING allows students to reflect on their knowledge of the concepts involved and construct a free-response answer, shown in Figure 1. It is important to note that this approach is not focused on solely getting individual responses nor is it focused on incorporating more technology into the classroom. This approach engages students in a complex cognitive task that causes the student to self-reflect as they compose their response.

These cognitive tasks can be thought of as a form of self-explanation, which has been shown in numerous studies to increase student learning gains (Chi, 2009). Importantly, SEEDING enables all students in the class to engage in this cognitive task, rather than just one student at a time, as is typically the case when a teacher asks a question in the classroom. We hypothesize that students using SEEDING to self-explain or articulate their beliefs about a subject will achieve learning gains similar to those seen in typical self-explanation scenarios.

Vocabulary List. Second language (SL) learners and students with low prior domain knowledge often struggle to articulate their explanations because they can't recall the right words. To aid these students in their self-explanation, SEEDING generates a vocabulary list. This list includes key content words extracted from the question's reference answer as well as various foils to mitigate the possibility of providing too strong of cues to the answer. Key content words are determined by their mutual information with the other questions and reference answers that the teacher saved in the same folder as the question being asked. The distractor words include key content words from those same related questions and their reference answers, WordNet's (e.g., WordNet is a freely available, machine-readable, lexical database for English available at: http://http://wordnet.princeton.edu) antonyms of the other words in the vocabulary list, and WordNet coordinate terms.

All the words in the vocabulary list are lemmatized, to extract the root. Repeated lemmas and words in the question, which the student can already see, are removed from the list. Only the most relevant distractors, those whose mutual information with the reference answer was the highest, are kept. Through teacher use, we empirically determined that ten words was the best number to keep. Finally, SEEDING presents the alphabetized list to the teacher, who is free to add or remove words from the vocabulary box and to send the list to any individual or to all logged in students. Ultimately, SEEDING aims to cognitively engage all students in self-explanation as they are constructing their responses and the vocabulary list can help by keeping SL learners and students with low prior knowledge engaged in the self-explanation process.

4 Enhanced Discussion

As students respond to a question, SEEDING performs analysis and provides teachers real-time feedback on the students' understanding. This is accomplished with system components such as a word cloud, clustering, and immediate presentation of individual student responses. The word cloud is updated in real-time to reveal the concepts students are focusing on in their responses. Clustering provides the teacher with representative responses from up to four primary groups of similar student responses, The presentation of individual student responses allows the teacher to check in on struggling students. Teachers can utilize all of this real-time feedback to evaluate whether or not the classroom understanding is headed in the direction they intend and decide what course corrections are necessary to clear up any issues or misconceptions.

Word Cloud. As students are constructing their free-response answers, SEEDING presents the teacher with a word cloud. A word cloud is a presentation of words that populates itself with frequently used content words. In this case, the word cloud is populated with words extracted from all of the student responses. A word is only presented to the teacher if it is used by more than one student. The more students that use a content word, the larger it will appear in the word cloud. The word cloud allows teachers to begin to assess the class' understanding before students submit their final responses.

Clustering. After students have submitted their responses to the teacher, SEEDING automatically clusters the responses in up to four groups based on

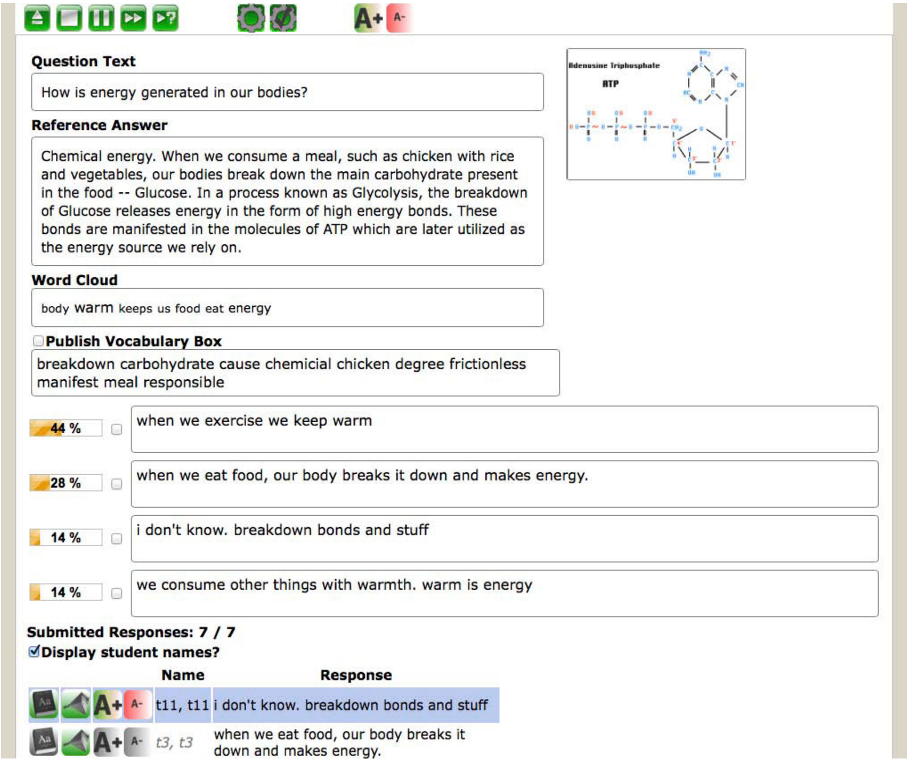

Fig. 1. Teacher control dashboard. Teachers view the word cloud, cluster representatives, and student responses.

semantic similarity. SEEDING will then present the teacher with a representative response for each cluster along with the percentage of student responses belonging to that particular cluster as shown below in figure 1. A cluster's representative is the student response that is the most representative of all of the responses in that cluster. The teacher has the option to share any or all of the cluster representatives with the class. Clustering and representative processing is hypothesized to facilitate meaningful classroom discussion because the teacher is presented with a sample of responses that represents the diverse views of the classroom. In addition, the teacher could address misconceptions in cluster representatives, ask the students to edit and resubmit their responses, and re-cluster the student responses.

To cluster student responses, we need an understanding of each student's response and its entailment relationship to the question's reference answer. We do not simply want to label responses as correct or incorrect. Instead if a response is not correct, we want to identify where the student's response is different from the reference answer and in what way it is different. To achieve this level of semantic analysis, SEEDING decomposes the question, its reference answer, and all the responses into their fine-grained semantic facets following (Nielsen et. al, 2009). An analysis of all of these semantic facets is used to generate the feature vectors used by the clustering algorithm, as discussed below.

Feature vectors are comprised of four sets of features, each of which is assigned a total weighting or importance. The sum of the weights over the four sets of features is 1.0. The first set of features is based on the subset of semantic facets found in the reference answer that are not also found in the question. These features were given a weight of 0.45. The second set of features, which has a weight of 0.225, is based on the remaining facets found in the reference answer (i.e., those facets that also existed in the question). The third set of features, with a weight of 0.1, is based on the facets found only in the question. The final set of features, comprising the remaining weight of 0.225, is based on any additional facets that occur in multiple student responses. In future work, the weights of each set of features will be learned based on training data. In the present work, facets from the reference answer were given most of the weight (just over 2/3 of the total weight), since those are the primary semantics of interest. Since it is easy for a student to just repeat words from the question, related facets were given less weight. Student responses are converted into feature vectors according to which facets in these four groups is entail by the response. These vectors are then used in the clustering process.

SEEDING automatically initiates the clustering when the percentage of students that have responded surpasses a threshold.[1] However, teachers have the option to cluster the responses much earlier, if desired, and are free to re-cluster the responses at any time, if they want to account for more complete information. Each time the teacher clusters responses, the system recomputes the feature vectors for any student response that has changed.

At the core of SEEDING's clustering is the k-means algorithm, shown in the equation below. Given a set of student responses, the goal is to find the assignment of responses, x_j, to k clusters, S = $\{S_1, S_2, ...S_k\}$, that minimizes the sum of the squared distances between the response vectors, x_j, and their associated (nearest) cluster centroid, μ_i.

Once all student responses have been converted into feature vectors. Four randomly selected student response vectors are assigned as the initial cluster centroids. We iterate over each student response vector, calculate its distance from each cluster centroid, and assign the response to the cluster whose centroid is closest. After each iteration, the cluster's centroid is recalculated by averaging the response vectors assigned to it. These two steps, assigning responses to the closest cluster and recomputing the cluster centroids, are repeated for 10 iterations or until convergence, when the clusters stop changing.

Following the clustering, representative responses are selected for each cluster. These representatives are presented to the teacher, who can use them to lead a classroom discussion focused on the main beliefs expressed by students. For each cluster, the response whose vector is determined to be closest to the cluster's centroid is selected as the cluster representative.

[1] In the present work, teacher feedback indicated that 50% was a reasonable threshold to present the teacher with cluster representatives.

These cluster representatives provide the teacher with a good sense of the student conceptions in the classroom. The teacher projects the representative responses onto the classroom display and engages the students in a discussion based on the various beliefs exemplified. Unlike clickers, which only allow teachers to guess a priori when writing the distractors what the misconceptions might be, SEEDING's Enhanced Discussion can directly target the beliefs held by the teacher's students. Unlike typical classroom discussions, which engage and address the perspective of only a single student at a time, SEEDING's dialogue is grounded by the diverse beliefs held in the teacher's classroom.

5 INquiry Generation

The question generation component of the SEEDING project is designed to expand the classroom discussion to a view of the topic as explored in the wider world, and to inspire teachers to think of science as a verb, not a noun. That is, science is not a static body of factual knowledge but a process of exploration, discovery, and peer review. The question generation component itself is being introduced in phases which represent different approaches to question generation. Phase I involves questions from the QtA Questioning the Author (Beck, 2001) framework, which has also been included in teacher training. Phase II utilizes questions extracted from the web. Phase III requires the development of a knowledge base, from which conceptual questions can be generated.

The Phase I QtA component takes all student responses as input, as well as the teacher question and reference answer. Common ideas are identified in the student responses by means of word frequency counts. Meanwhile, the teacher question is analyzed to see if a concept can be extracted. For each noun in the teacher question, mutual information is calculated between these nouns and the question category extracted from within the SEEDING system. The highest scoring noun is selected as the concept, with preceding nouns and prepending adjectives, as in *kinetic energy.* There are over 100 QtA question stems which are divided into subsets for random selection based on whether the teacher question referenced a lab or experiment, whether a teacher question concept or student common idea was identified, or one of the remaining question stems. Sample stems include:

– Can you think of another experiment we could do which would teach us more about *concept*? If you were explaining *concept* to a younger person, what other knowledge would they need to understand your explanation?
– Many of you mentioned *common idea.* Does anyone disagree?
– After reading the responses on the screen, what would you change about your response, and why? If you would not change your response, why is yours better?

The questions extracted from the web in Phase II utilize the teacher question and reference answer in the web search. These texts are tokenized and tagged by the Stanford taggers, and stop words are removed. Words from this group with the desired parts of speech (nouns, verbs, adjectives) are extracted as keywords. These keywords are sent to a Google custom search engine to retrieve relevant urls. A web crawler then traverses these urls, and the links from those pages, to extract all questions from the pages it crawls. Questions are rated according to the frequency of the keywords, and the top ranking questions are sent to be displayed. For example, the teacher question How is work turned into mechanical energy? results in the keywords: work, turned, mechanical, and energy. The top retrieved questions are:

– What devices convert mechanical energy to heat energy?
– How can mechanical energy be converted to heat energy?

Note that these questions extend the discussion beyond the original teacher question to more application and conceptual questions. The urls from which the questions were retrieved are also provided to the teacher.

6 Discussion

As of the spring 2014 phase, over 1200 students are using SEEDING in their classrooms. We collect feedback from the teachers and make changes to the system accordingly. As a result, new ways to enhance the classroom learning environment are still being developed.

Evaluation in Progress. To evaluate the effectiveness of the Comprehension Seeding system compared to traditional and clicker classrooms, we are conducting a yearlong pilot study within sixth grade science classrooms. We are analyzing the effect of the SEEDING system use on student learning, in addition to learning more about SEEDING adoption, use and integration into teacher practices. With respect to teacher adoption and use, we have collected a substantial amount of data from the teachers starting with the participatory design process and following all the way through system deployment and use. This data consists of informal interviews with teachers, short surveys, frequent email follow-ups, and discussions during researcher and support team visits. To date, the teachers have been very forthcoming with their system design needs, desires, issues, and potential barriers to use. This information has contributed greatly to our ability to make the system and interface "teacher friendly." We also collected a very substantial amount of observation and system log data related to teachers' use of the system in practice. This data helps us to make sense of how the teachers are integrating the system into their practice. As a specific example, we would hope that the teachers use the system to gather class-level formative feedback that will help them lead a rich follow-up discussion. Observation and logs can tell us if teachers are asking follow-up questions to the initiating questions, how long those questions are open for student responses, and whether or not the teacher pauses the question during student response (perhaps to discuss or clarify). In this way we are able to identify any specific pedagogical needs that the teacher may have in order to fully integrate the system into their classroom practice.

Teachers' (and students') feedback on the system has been overwhelmingly positive. The teachers' especially appreciate the fact that all students can individually respond to a question, and that student responses can be displayed for class discussion. Students enjoy expressing their own thoughts, and become very excited when their responses are displayed as one of the cluster representatives.

We are in the process of collecting student assessment data to investigate the effect of the system on student learning. We have structured a within-teacher research design in order to control for teacher effects. Any given teacher in our research is teaching one or more class sections using the system, and other sections using clickers (multiple choice only) or no technology support. We have designed our own assessments of students' deep learning in four science units: Atoms & Elements, Particulate Model of Matter, Force & Motion, and Energy. These assessments consist of both open ended and multiple choice items that span a range of cognitive depth. Each class section (SEEDING, clicker, or no technology support) responds to each unit test pre and post instruction for that unit. The students also respond to a year long pre and post test which encompass all of these topics. This data collection and the scoring of the student responses is ongoing.

Rather than collecting this assessment data with paper tests, we added a component to the SEEDING system specifically for this purpose. Using SEEDING, teachers specify what class and exactly how long an assessment should be. Once a teacher begins an assessment, students are redirected from the traditional interface and taken to an assessment page. This page allows students to submit answers through free-response, multiple choice, and canvas, where using a stylus, students can draw their responses to a question. While students are in assessments, they are free to navigate through all the questions in the assessment, edit their responses, or erase their drawings. Once the time for an assessment ends or the teacher decides to terminate the assessment, the students exit the assessment.

Vocabulary List. We plan to do future research that will lead to populating the vocabulary box with words more meaningful to SL learners. We are exploring using a large corpus as a filter to non-science related words. We do this by calculating co-occurrence relationships between science words. In addition, we are exploring extracting hypernyms from content words to provide a broader perspective of the given word.

Facet Cloud. To provide teachers with even more real-time information about student understanding as they construct their responses, we will explore a facet cloud. Similar to the word cloud, the facet-cloud will give teachers an indication of how many students expressed each semantic facet. This will allow teachers to see the semantic relationships students make as they type out their responses. For example, if a teacher asks *Is a proton positive or negatively charged?* as students are responding, the facet-cloud could present facets such as: (proton, neutral), (atom, positive), etc. Teachers can use this feedback to guide the classroom discussion accordingly.

7 Conclusion

It is expected that combining the scientifically-grounded educational support technology and methods in Comprehension Seeding will result in learning gains that could exceed the one sigma gain found in the best current tutoring systems as well as the more modest gains associated with effective implementation of clicker systems. From a cost-benefit perspective, Comprehension SEEDING has the potential to inexpensively provide a practical, focused, nearly individualized, adaptive, scientifically based solution. Furthermore, this solution is not tied to one specific inquiry-based pedagogy or to science education, but rather has the potential for significant positive impact across many areas in education. We are currently conducting a study involving approximately 1250 students to assess the impact of Comprehension SEEDING in the classroom.

Acknowledgements. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A120808 to UNT. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

1. Beck, I., Margaret, M.: Inviting students into the pursuit of meaning. Educational Psychology Review 13(3), 225–241 (2001)
2. Bunce, D.M., VandenPlas, J.R., Havanki, K.: Comparing the effectiveness of student achievement of a student response system versus online WebCT quizzes. Journal of Chemisty Education 83(3), 488–493 (2006)
3. Carnaghan, C., Webb, A.: Investigating the Effects of Group Response Systems on Student Satisfaction, Learning, and Engagement in Accounting Education. Issues in Accounting Education 22(3), 391–409 (2007)
4. Chi, M.T.H.: Active-Constructive-Interactive: A Conceptual Framework for Differentiating Learning Activities. Topics in Cognitive Science 1(1), 73–105 (2009)
5. Duggan, P.M., Palmer, E., Devitt, P.: Electronic voting to encourage interactive lectures: A randomised trial. BMC Medical Education 7(25) (2007)
6. Duncan, D.: Clickers: A New Teaching Aid with Exceptional Promise. Astronomy Education Review 5(1), 70 (2006)
7. Fies, C., Marshall, J.: Classroom Response Systems: A Review of the Literature. Journal of Science Education and Technology 15(1), 101–109 (2006)
8. Herreid, C.F.: Clicker Cases: Introducing Case Study Teaching Into Large Classrooms (2006)
9. Keller, C., Finkelstein, N., Perkins, K., Pollock, S., Turpen, C., Dubson, M., Hsu, L., et al.: Research-based Practices For Effective Clicker Use. In: AIP Conference Proceedings, pp. 128–131 (2007)
10. Larson, B.: Classroom discussion: A method of instruction and a curriculum outcome. Teaching and Teacher Education 16(5-6), 661–677 (2000)

11. Nielsen, R.D., Ward, W., Martin, J.H.: Automatic Generation of Fine-Grained Representations of Learner Response Semantics. In: Woolf, B.P., Aïmeur, E., Nkambou, R., Lajoie, S. (eds.) ITS 2008. LNCS, vol. 5091, pp. 173–183. Springer, Heidelberg (2008)
12. Penuel, W.R., Boscardin, C.K., Masyn, K., Crawford, V.M.: Teaching with student response systems in elementary and secondary education settings: A survey study. Educational Technology Research and Development 55(4), 315–346 (2006)
13. Siau, K., Nah, F.F., Siau, K., Sheng, H., Nah, F.F.: Use of a Classroom Response System to Enhance Classroom Interactivity (2006)