

---

# Assessment of KNOWLA: Knowledge Assembly for Learning and Assessment

**Meredith Thompson**

Harvard University  
13 Appian Way  
Cambridge MA 02138  
Meredith\_thompson@  
gse.harvard.edu

**Eric Braude**

MET College 755 Comm. Ave.  
Boston University  
Boston MA 02215  
ebraude@bu.edu

**Christopher D. Canfield**

MET College 755 Comm. Ave.  
Boston University  
Boston MA 02215  
cdcanfie@bu.edu

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/author(s).  
L@S 2015, Mar 14-18, 2015, Vancouver, BC, Canada  
ACM 978-1-4503-3411-2/15/03.  
<http://dx.doi.org/10.1145/2724660.2728673>.

**Jay Halfond**

MET College 755 Comm. Ave.  
Boston University  
Boston MA 02215  
jhalfond@bu.edu

**Aparajita Sengupta**

MET College 755 Comm. Ave.  
Boston University  
Boston MA 02215  
asgupta@bu.edu

**Abstract**

The assessment of learning in large online courses such as Massive Online Open Courses, or MOOCs, requires tools that are valid, reliable, and can be automatically administered and scored. We have developed and assessed a tool called Knowledge Assembly for Learning and Assessment, or KNOWLA. The tool measures a student's knowledge in a particular subject by having her assemble a set of scrambled phrases into a logical order. Initial testing indicates that KNOWLA is reliable, and can be used to measure learning gains. KNOWLA also shows promise as a learning tool.

**Author Keywords**

Assessment; learning aids; scrambled text; automatic scoring; college-level; massive online open source courses

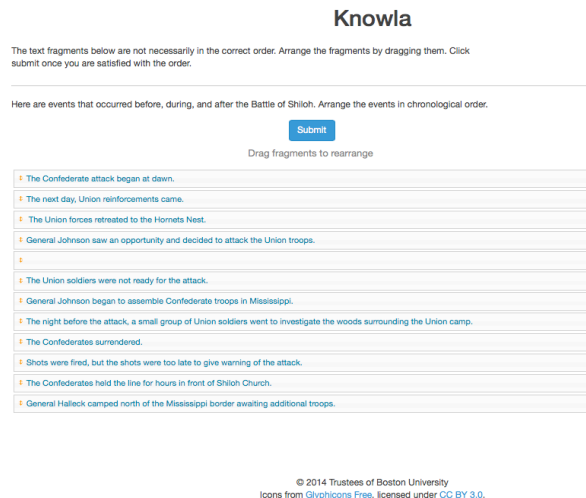
**ACM Classification Keywords**

Student assessment, adult education, measurement, arts and humanities

**Introduction**

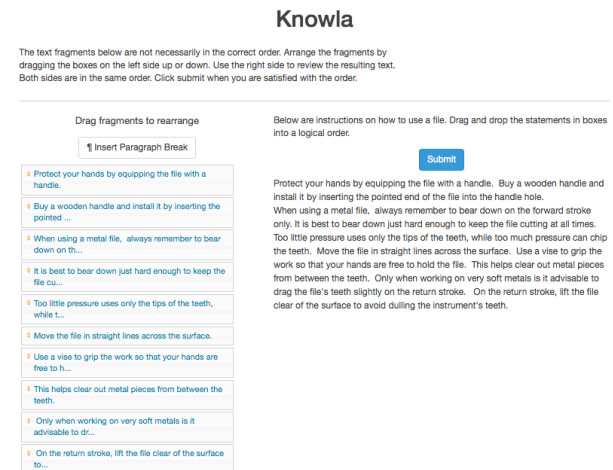
The advent of online education, and the hope that digital technology can extend the reach of higher education, present a new challenge for finding assessment tools beyond traditional multiple-choice questions that can also be automatically scored.

Although many advances have been made in the areas of intelligent tutoring systems [1] and digital dashboard displays [2], additional assessment and learning strategies are needed. This study describes a dynamic tool called Knowledge Assembly for Learning and Assessment or *KNOWLA*, originated by Braude and implemented by Braude and Canfield. In a *KNOWLA* assessment, the test-taker demonstrates knowledge of a topic by unscrambling a series of sentences, phrases, or fragments via an online interface. *KNOWLA* allows instructors to (1) scramble a text or procedure (such as a text passage, science experiment, or computer program), (2) develop rubrics, (3) provide each student with a different scrambled order that can be dragged and dropped, (4) automatically score the results based on the rubrics, and (5) automatically provide comments.



**Figure 1:** A full screen *KNOWLA* layout.

Instructors upload a plain text file in which a delimiter indicates how it should be fragmented. Instructors can select one of two possible *KNOWLA* displays for students. In both configurations, instructions appear at the top of the window, and segments of text that may be phrases or sentences are displayed in boxes that can be dragged and dropped into the desired order. In the full screen version shown in Figure 1, the boxes of text span the entire screen. In the side-by-side configuration shown in Figure 2, the moveable boxes appear on the left side of the screen and the statements appear as unboxed text in the same order on the right side of the screen.



**Figure 2:** A side-by-side *KNOWLA* layout.

### **KNOWLA Grading Scheme**

*KNOWLA* tests are scored according to the level of importance that the instructor associates with the order of fragments, and in such a way that a perfectly

ordered answer receives 100%. The default (or standard) scoring gives equal weight to every correctly ordered (though not necessarily consecutive) pair. Suppose, for example, that the correct response of a 5-fragment test is 1-2-3-4-5. A 1-4-2-3-5 response would receive a score of 80% when default scoring is used since two of the ten correct pair orderings (1-2, 1-3, 1-4, 1-5, 2-3, 2-4, 2-5, 3-4, 3-5, and 4-5) are missing. The tests discussed in this paper were graded with the default rubric. The instructor can impose two kinds of rubrics in addition (other kinds will be considered at a later time). An *order* rubric allows the instructor to give the weight of an order. A *consecutivity* rubric allows the instructor to specially value pairs that occur consecutively in the response. These types of rubrics will be discussed in detail elsewhere.

### **Study**

KNOWLA is designed to be one tool to assess learning, and a KNOWLA question can be used alone either as a formative tool such as a quiz or as one question on a larger summative tool such as a final test or exam [3]. This study examines whether KNOWLA can show evidence of short-term learning using a pre-test/ post-test design. In particular, we want to answer the following questions: How does studying a passage influence performance on a KNOWLA assessment? Does performance improve on KNOWLA assessment if the passage is not studied?

### *Method*

Fifty graduate students in the School of Education at Boston University participated in this sub study during the course "Introduction to Educational Research Methods". After written consent was obtained, students took two KNOWLA tests and answered multiple

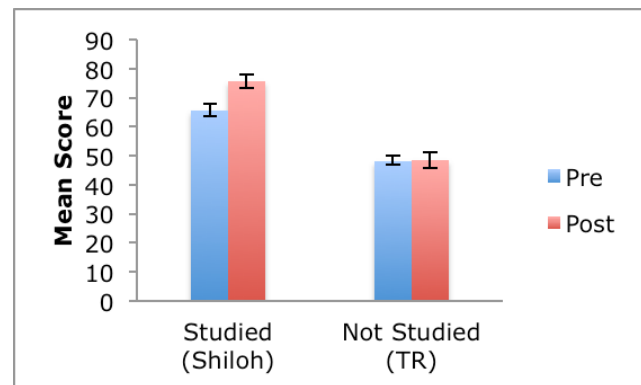
questions about two topics, an inauguration speech by Theodore Roosevelt, and a historical account of the Battle of Shiloh in the Civil War. The researcher Thompson then gave a short presentation reviewing the passage about the Battle of Shiloh using four PowerPoint slides. Students were then asked to take the same two KNOWLA tests and answer the same multiple-choice questions as in the pre-test. The pre-test post test study design with two forms of assessment (a set of multiple choice questions and a KNOWLA assessment) for each topic allowed us to identify whether learning happened as a result of the teaching event as indicated by an increase in multiple choice question scores for the topic studied. The study design also allowed us to identify whether a KNOWLA assessment could capture learning gains. Learning gains are indicated by an increase in the correct ordering of the statements in the KNOWLA test for the passage that was studied.

### *Results*

The results from the online survey and KNOWLA tests were downloaded and queried together in Excel. Data were imported into SPSS 22.0 for windows for the analysis. The mean KNOWLA scores on the post-test of the Battle of Shiloh ( $M=75.85$ ,  $SD=11.74$ ,  $n=43$ ) were higher than the mean pre KNOWLA tests ( $M=65.54$ ,  $SD = 15.86$ ,  $n=43$ ,  $p<0.001$ ). For the passage that was not studied, participants did equally well on the pre ( $M=48.22$ ,  $SD=15.47$ ) and the post ( $M=48.50$ ,  $SD=18.23$ ,  $p=0.902$ ) for the Theodore Roosevelt (TR) passage that was not studied. The effect size was calculated using Cohen's  $d$  for paired t-tests [1]. The results was a Cohen's  $d$  of .55, (95% CI=0.115 to 0.988,  $v=0.0497$ ), which is considered to be a large

effect [4]. The results of the t-test are summarized in Figure 1. Error bars in the figure show standard error.

A series of multiple choice questions were also asked about the Shiloh and TR topics for a point of comparison. Participants did slightly better on the post-test multiple choice questions about Shiloh ( $M=1.63$ ,  $SD=1.02$ ) than on the pre-test ( $M=1.30$ ,  $SD=1.01$ ), but the difference was not statistically significant ( $p=0.12$ ). Participants did worse on the post questions about Theodore Roosevelt ( $M=2.72$ ,  $SD=1.07$ ) than they did on the pre questions about the Theodore Roosevelt topic ( $M=2.89$ ,  $SD=0.85$ ). That difference was statistically significant ( $p=0.038$ ). The slight gain for the Shiloh questions suggests that learning occurred as a result of the short lesson. The statistically significant decrease for the Theodore Roosevelt questions cannot be explained with the information gathered through the current study and will have to be explored through additional research.



**Figure 1:** Paired t-test results from the studied (Shiloh) and non-studied (TR) passages

Overall, the participants made significant gains on the KNOWLA test of the passage that was studied, and they scored almost exactly the same on the pre- and post-KNOWLA test for the passage that was not studied.

## Discussion

A principal goal for the study was to examine whether KNOWLA can be used as a reliable tool to measure learning. Students completed KNOWLA assessments on two passages, learned about one passage, and then completed two additional sets of multiple choice and KNOWLA questions. Since participants scored higher on the post-test for the studied passage than the pre-test and maintained the same average scores for the non-studied passage, we can conclude that KNOWLA may be useful as a reliable and valid measure of learning about a specific topic.

This study does not provide a clear picture of what specific attributes a KNOWLA assessment does test; however, scrambled text has been used as a measure of academic literacy [6], reading comprehension [7, 8], and writing skill [8] in other studies. We are currently investigating the link between KNOWLA scores and those skills.

Some of the participants noted that they learned more about the topic from doing the test than had they simply read the paragraph quickly. In fact, scrambled text has been used for self-assessment in medical education [9]. Students who engaged more deeply in scrambled reading material are more likely to learn from it than from even well-organized text [10]. The KNOWLA format may become a useful learning tool, as well as a means of assessment.

## Conclusion

These preliminary findings support KNOWLA as a reliable tool that can be used to measure learning. Future research could explore other constructs that may provide a better match for the type of critical thinking captured by a KNOWLA test. Since KNOWLA tests require students to engage deeply with the material, KNOWLA also has potential as an engaging tool for learning and reviewing content.

## Acknowledgements

We thank the Digital Learning Institute (DLI) at Boston University for project funding and Chrysanthos Dellarocas and Romy Ruukel for their support. We thank Bruce Fraser, Laura Jimenez, Matthew Parfitt, Natalie McKnight, Gillian Pierce, John McGrath, Davida Pines, Tejvir Sekhon, the Behavioral Lab at the School of Management, and the study participants.

## References

- [1] Bajzek, D., Brooks, J., Jerome, W., Lovett, M., Rinderle, J., Rule, G. & Thille, C. (2008). *Assessment and Instruction: Two Sides of the Same Coin*. In G. Richards (Ed.), Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2008 (pp. 560-565). Chesapeake, VA: AACE.
- [2] Psootka, J., Massey, L.D., and Mutter, S. A. eds. *Intelligent tutoring systems: Lessons learned*. Psychology Press, 1988.
- [3] Gronlund, N.E. *How to make achievement tests and assessments*. Needham Heights, MA, USA, 1993.
- [4] Wilson, D.B. *Practice Meta-analysis Effect Size Calculator*. 2014. [cited 2014 Dec 17]; Available from: <http://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-SMD25.php>.
- [5] Cohen, J. A power primer. *Psychological Bulletin*, 1992. **112** (1): p. 155-159.
- [6] van der Silk, F. , Weideman, A., The refinement of a test of academic literacy. *Per Linguam*, 2005. **21**(2). p 23-35.
- [7] Gernabacher, M.A., Varner, K.R., Faust, M.E., Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 1990. 16(3): p. 430.
- [8] Kibby, M.W., Intersentential processes in reading comprehension. *Journal of Literacy Research*, 12(4) (1980) p 299-312.
- [9] Coirier, P., Favart, M., Chanquoy, L., Ordering and structuring ideas in text: From conceptual organization to linguistic formulation. *European Journal of Psychology of Education*, 2002. 17(2): p. 157-175.
- [10] Cairncross, R., Harden, R.. Preparation of scrambled text for use in self-assessment exercises. *Medical Education*, 1983. 17(4): p. 277-279.
- [11] Kintsch, W. Text comprehension, memory, and learning. *American Psychologist*. 1994. 49(4) p 294.