Comparing the efficiency of worked examples to erroneous examples, tutored problem solving, and untutored problem solving in e-learning environments

👤 **Lily Lou, Lauryn Jefferson, James Wei, Malik Richardson**
Human-Computer Interaction Institute, Carnegie Mellon University

SUMMARY

A major and recurring question for teachers and developers of instructional software is how much guidance or assistance they should provide in order to lead to the best learning outcomes for students. This is known as the 'assistance dilemma' (Koedinger & Aleven, 2007).

METHODS

- Web-based stoichiometry-learning environment was implemented with the different learning conditions
- Student performance per condition was determined through pretest, posttest and intervention problems
- Mental effort, time on task and post-questionnaire ratings were also measured

EXPERIMENT 1:

- 179 students from 10th and 11th grade
- Randomly assigned to (1) Worked Examples (n=39), (2) Erroneous Examples (n=43), (3) Tutored Problems (n=36), (4) Problems to Solve (n=37)
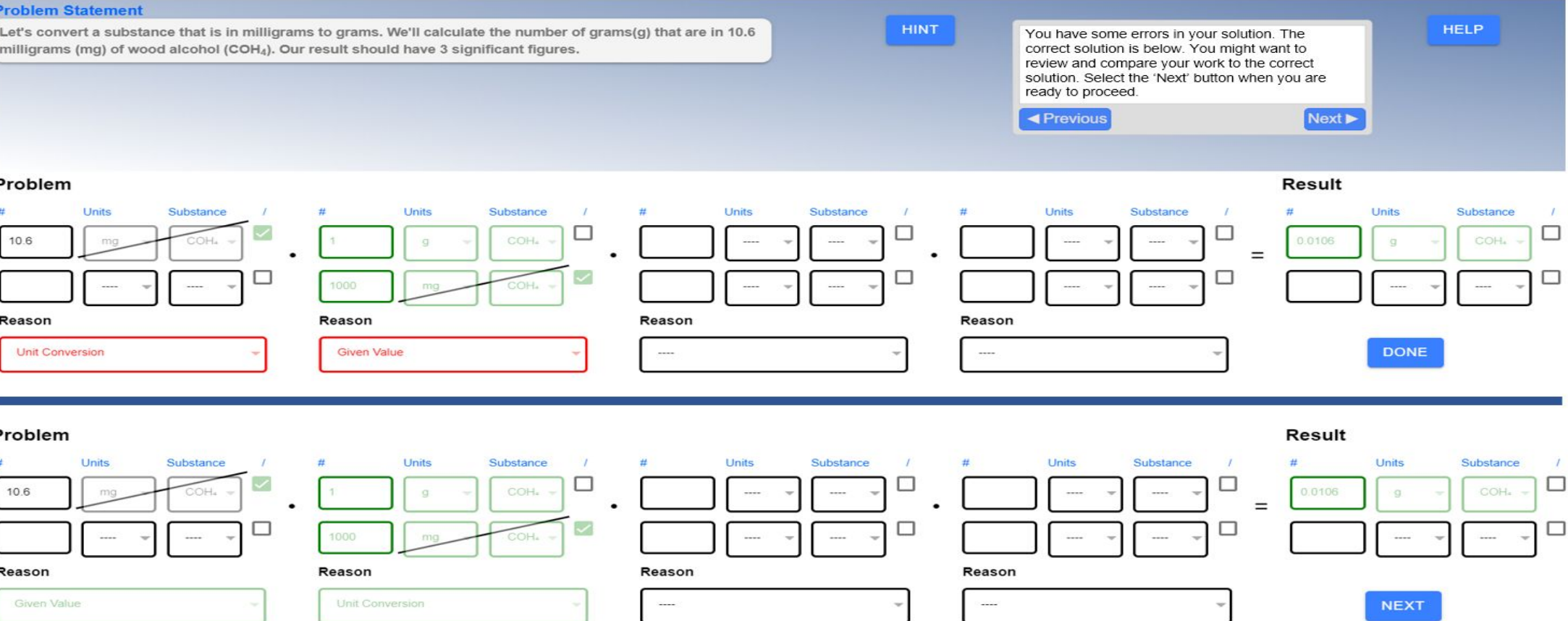- Error feedback was given as a worked example (Fig. 1)

Fig. 1. Worked Example from Experiment 1, with feedback indicating incorrect reasons selected and the correct worked example shown below the student's work.

EXPERIMENT 2:

- 131 students from 10th and 11th grade
- Randomly assigned to (1) Worked Examples (n=29), (2) Erroneous Examples (n=28), (3) Tutored Problems (n=27), (4) Problems to Solve (n=32)
- Error feedback highlighted steps that were correct (green) or incorrect (red) (Fig. 2)

Fig. 2. Tutored Problems Example from Experiment 2, with feedback indicating incorrect reasons selected (red) and correct reasons selected (green).

For students, worked example study is shown to be **~46%-69% more efficient** than other instructional approaches.
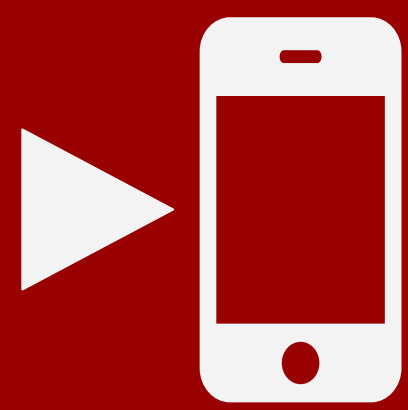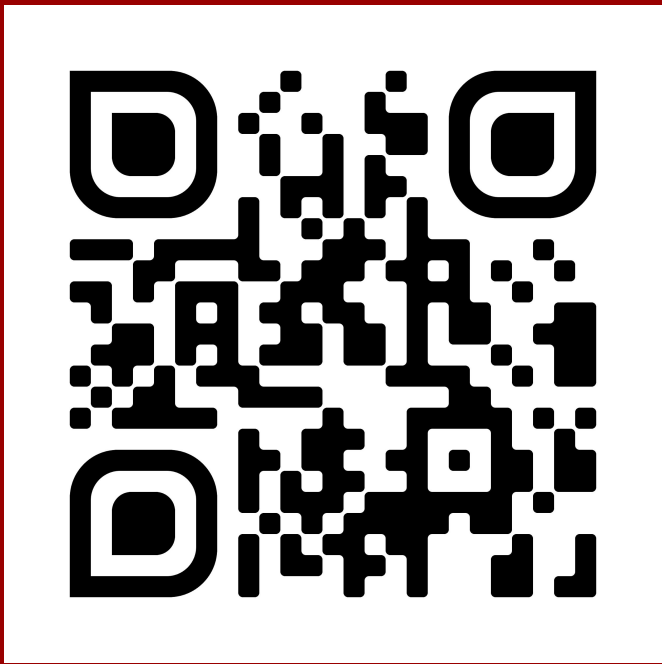
Take a picture to **download** the **full paper**

RESULTS

- There were no differences in learning across conditions
- Worked example study was much more efficient in terms of time and effort spend on the intervention problems.

Table 1
Performance, mental effort, time on task, and post-questionnaire ratings per condition in Experiment 1.

	Condition:			
	WE (n = 39)	ErrEx (n = 36)	TPS (n = 43)	PS (n = 37)
Pretest (max = 101)	48.6 (12.8)	48.8 (15.4)	49.4 (13.5)	46.3 (14.3)
Posttest (max = 101)	68.5 (17.3)	68.3 (18.4)	71.1 (13.4)	66.4 (17.1)
Embedded test (max. = 122)	89.4 (23.7)	88.3 (27.0)	95.3 (23.3)	84.8 (23.1)
Mental effort on intervention problems (1–9)	4.4 (1.8)	5.8 (1.4)	6.1 (1.7)	6.1 (1.3)
Time on intervention problems (min.)	19.8 (5.8)	37.2 (9.6)	62.4 (17.2)	52.1 (25.2)
Reflection time on feedback (min.)	1.7 (1.1)	4.3 (2.6)	1.3 (1.0)	6.5 (3.9)
Posttest confidence (correct out of 8; N = 135)	4.9 (2.3)	4.7 (1.9)	3.9 (2.4)	4.3 (2.0)
Liked materials (1–5; N = 135)	2.7 (1.2)	2.7 (1.3)	2.6 (1.3)	3.3 (1.1)
Want to work again with materials (1–5; N = 135)	2.2 (1.0)	2.6 (1.2)	2.3 (1.2)	2.8 (1.1)

Note. Time on task concerns only the intervention problems which differed among conditions; it does not include the time spent on the pre-questionnaire, pretest, instruction videos, effort ratings, embedded test problems, post-questionnaire, and posttest.

Table 2
Performance, mental effort, time on task, and post-questionnaire ratings per condition in Experiment 2.

	Condition			
	WE (n = 29)	ErrEx (n = 28)	TPS (n = 27)	PS (n = 32)
Pretest (max. = 101)	48.7 (17.6)	47.5 (20.3)	41.9 (16.8)	45.3 (16.3)
Posttest (max. = 101)	68.2 (18.2)	65.7 (23.1)	67.8 (20.0)	69.9 (19.2)
Embedded test (max. = 122)	92.2 (25.0)	79.8 (33.3)	85.3 (30.9)	80.7 (31.1)
Mental effort on intervention problems (1–9)	4.9 (1.4)	5.3 (1.7)	6.7 (1.3)	6.3 (1.3)
Time on intervention problems (min.)	20.9 (5.5)	40.5 (11.3)	67.1 (18.9)	56.8 (11.8)
Reflection time on feedback (min.)	2.3 (1.6)	1.3 (.9)	1.5 (1.0)	1.8 (1.7)
Posttest confidence (correct out of 8; N = 135)	4.8 (2.1)	3.8 (2.0)	4.9 (1.7)	3.5 (2.2)
Liked materials (1–5; N = 135)	2.9 (1.2)	2.8 (1.2)	3.2 (1.2)	2.6 (1.3)
Want to work again with materials (1–5; N = 135)	2.8 (1.3)	2.8 (1.3)	2.9 (1.1)	2.2 (1.1)

Note. Time on task concerns only the intervention problems which differed among conditions; it does not include the time spent on the pre-questionnaire, pretest, instruction videos, effort ratings, embedded test problems, post-questionnaire, and posttest.

CONCLUSION & DISCUSSION

- This study is the first to directly compare the effects of four instructional methods that vary in the type and amount of assistance.
- A potential limitation of this study is that we did not assess whether learning benefits across the conditions would have varied on a delayed posttest.
- Future research should continue to compare the effects of varying degrees of instructional assistance in the classroom, given that direct comparisons are informative and multi-session experiments are scarce but much closer to the reality of educational practice.

References

Mclaren, B. M., Gog, T. V., Ganoe, C., Karabinos, M., & Yaron, D. (2016). The efficiency of worked examples compared to erroneous examples, tutored problem solving, and problem solving in computer-based learning environments. *Computers in Human Behavior, 55*, 87-99. doi:10.1016/j.chb.2015.08.038

Koedinger, K. R., & Aleven, V. (2007). Exploring the Assistance Dilemma in Experiments with Cognitive Tutors. *Educational Psychology Review, 19*(3), 239-264. doi:10.1007/s10648-007-9049-0

Acknowledgements

Carnegie Mellon University