

Project Goals

This four-year Level II Teaching Strand project, targeting early-stage Design and Development, has three goals:

- (a) develop automatically generated student reports (AutoRs) for threedimensional (3D) science assessments to assist middle-school teachers to notice, attend to, and interpret information in ongoing classroom teaching;
- (b) develop effective pedagogical content knowledge supports (PCKSs) to improve teachers' use of AutoRs to make effective decisions for instructional moves;
- (c) examine the effectiveness of AutoRs and PCKSs to support teachers' decision making and student 3D learning.

Theoretical Framework

We develop a theoretical model accounting for scaffolding teachers' instructional decision-making with AutoRs and PCKSs. In this model, AutoRs and PCKSs scaffolds teachers to make instructional decisions that influence students' 3D learning outcomes. AutoRs, as part of the 3D assessment, helps teachers interpret students' responses to 3D assessments. PCKSs are expected to improve teachers' transferable capacity because they can eventually improve teachers' 3D knowledge and beliefs so that teachers can apply PCK and relevant skills to new teaching scenarios.



Automatic scoring development

We select nine items about chemical reaction from NGSA and develop 3D-based analytic rubrics for each item. Each item has 5-10 dimensions. We use a pre-trained natural language processing model – BERT to do a binary text-classification task for automatic scoring. According to the analytic scores, we further divide student responses into four groups, which show the responses achievement on 3D.

| Gas filled balloons Alice did an experiment that caused four balloons to fill with gas, as shown in the figure to the right. Alice tested the flammability of each gas. She also measured the volume and mass of each gas to calculate the density. The tests and measures all occurred under the same conditions. The data is in the Table 1 below. | | | | | Student response: Gases A and D could be the same gas. Because they have same volume. | | | | | | |
|---|--|------------------------------------|---------------------------------------|---------------------|---|--|----|------------|----------|---------|--|
| | | | | | | | | Dimensions | Elements | | |
| | | | | | | | | | | | |
| | | | | | | | p1 | el atras | - | SEP+CCC | Student supports a claim by referring to a |
| | | b | | | | Gas A and Gase D is the same in the tabl | | | | | |
| | | | | | | columns). | | | | | |
| | Table 1. Data o | ata of four gases in the balloons. | | SEP+CCC | Student supports a claim by referring to a | | | | | | |
| | Sample | Flammability | Density | Volume | | A and Gas D is the same in the table. (co | | | | | |
| | Gas A | Yes | 0.089 g/L | 180 cm ³ | | columns) | | | | | |
| | Gas B | No | 1.422 g/L | 270 cm ³ | DCI | Student indicates flammability is one of | | | | | |
| | Gas C | No | 1.981 g/L | 35 cm ³ | | Student indicates frammability is one of | | | | | |
| | Gas D | Yes | 0.089 g/L | 269 cm ³ | | identify substances. | | | | | |
| Which, if any, of t | the gases listed in the data table could be the same? Using information from the table, explain your answer. | | | vour answer. | DCI | Student indicate density is one of the cha | | | | | |
| | | | · · · · · · · · · · · · · · · · · · · | | | substances | | | | | |

Making: The Potential of Automatically Scored Threedimensional Assessment

Xiaoming Zhai¹, Joe Krajak², Yue Yin³, Christopher Harris⁴, Xinyu He¹, Ehsan Latif¹, Peng He², Namsoo Shin², David McKinney⁴ ¹University of Georgia; ²Michigan State University; ³University of Illinois Chicago; ⁴WestEd

AutoR Design

Inspired by cognitive load theory, we propose a conceptual framework for teacher-centered AutoRs design, including cognitive demands and human-centered design support. Cognitive demands. This dimension examines the amount and level of integration of the information in the report. Three characteristics are recognized to count the amount and the level of the information: the content presented in the report, the synthesis level of presented information, and the depth of data mining.

Synthesis

GNITIVE DEN Human-centered design support. This dimension examines the design features that can assist teachers in interpreting and using the information. We identify two sub-dimensions: user functionality and information presentation.

To provide grouping information, we develop a grouping rubric based on students responses achievement on 3D.

We plan to present three types of information on screen: a) grouping information,

b) instructional strategies suggestions; and

c) individual task performance. Then, we have teachers involved to provide feedback of the AutoR and revise the AutoR design in an iterative loop.

| | Class 1 🔻 | Lesson 3 | Gas Filled Balloons Accuracy: 0.902 | 26 |
|--|----------------------------|-----------------------------------|---|--|
| | DIAGNOSTIC | Group by Similarity 🔻 | | Mary |
| Home Questions Submissions | Group Blue Jim Matt Sue | Group Green | Group Orange Need support in understanding characteristic properties 30 % of your students have identified data patterns, but they need your support in understanding density and flammability as the characteristic properties. | Gas b and c are both no and there density are al would look at the flamm Also gas a and d both a and have density the sa |
| Report | Group Orange | Group Red Tony Tim <u>Stev</u> | Teaching Strategies | Historical Perform Tube Su Group Blue |
| Class Settings | Mar Dan Ted | Mao Lisa Ann Peng Rich Lilly | Activity 1: three samples of the same metaleq., Al or Fe) but different volumes and masses. Activity 2: Students have three samples of metals (e.g., Al and Fe) with the same volume or mass. | Group Green Group Orange Group Red |
| | | | Second-Hand Experiences: Students analyze data using the presented data table. Activity 1: Density with Five balloons containing different gases Activity 2: Solubility with Five balloons containing different gases Simulation: Student use a PhET simulation related to the density and choose mystery as the task. | Student Comm |
| | | | PhET simulation on five blocks and identifying a mystery block based on the density information | Quick Feedback |

| the same density even though they d | lon't have the | | % of | Testing Algorithm- Tensor Flow | | |
|---------------------------------------|----------------|----------|----------|-----------------------------------|----------|-------|
| | Score | | students | (Trans | sformer- | Bert |
| could be the same substance. | 1 | Question | correct- | Based) | | |
| data pattern that the flammability of | 0 | | human | | | |
| e. (comparisons of data in different | | c01Gas | scoring | Aspect | Accuracy | Kappa |
| data pattern that the density of Gas | 1 | filled | 74.30% | 1 | 95.98 | 0.885 |
| nparisons of data in different | | balloons | 56.90% | 2 | 96.77 | 0.934 |
| he characteristic properties to | 0 | (034.02- | 65.90% | 3 | 95.54 | 0.902 |
| le characteristic properties to | U | c01) | 10.10% | 4 | 95.54 | 0.783 |
| acteristic properties to identify | 0 | | 11.30% | 5 | 96.76 | 0.837 |
| | | | • | • | • | • |

| Group | Activeued | 1 | | |
|-------|--|--|--|--|
| A | Students have achieved in 30. • DOI indeptioning the characteristic properties in non-characteristic properties of satisfamos • MOP producing and interpreting data • DOI: identifying patterns in data. | Students have a te move forward | | |
| B | Students have achieved in DO. • DOI understanding the characteristic properties us non-characteristic properties of substances. | Students mend in • SCPL analysing at • CCC sideot/ying) | | |
| c | Students have achieved in SEP and CCC. • SEP: unalgoing and incorporing date and • CED: identifying patterns in data | Students need at • OC understand characteristic pro- | | |
| 0 | Students have achieved in N/A. | Students need a • OC understand characteristic pr • SCP analyzing of • CCC interst/ving) | | |



| Algorithm- | | | | | |
|------------|-------|--|--|--|--|
| or Flov | v | | | | |
| rmer- Bert | | | | | |
| sed) | | | | | |
| | | | | | |
| curacy | Kappa | | | | |
| 95.98 | 0.885 | | | | |
| 96.77 | 0.934 | | | | |
| 95.54 | 0.902 | | | | |
| 95.54 | 0.783 | | | | |

Teacher Interpretation of AutoR

- Knowledge activation. The original AutoRs activate teachers' curiosity about students' learning. Teachers connected the information received with their experience in class (e.g., a specific student).
- Contextualization. Teacher abstract the contextual information to understand the situation and process. For example, Teachers A and B paid attention to the scores of some students who were highlighted in the diagram of responding time and showed their concerns in figuring out the learning condition of the students.
- Cross-validating information for reasoning. In reading the individual scores table, teachers inferred the highlighted students' thinking who were identified before.
- Prioritizing information for reasoning. Teacher C paid attention mainly to the individual scores table and made decisions directly based on the scores. However, teacher A spent most of the time reviewing the group descriptions.



Instructional Strategies for Using AutoRs

A second challenge for teachers is how to effectively transform student performance on 3D assessments into effective instruction. Effective instructional moves require not only the ability to interpret assessment results but also actionable knowledge (Bennett, 2018). If teachers have limited PCK to transform their interpretation of student performance into meaningful 3D instructional activities, 3D assessments might still end up with a limited impact. Therefore, essential PCKSs such as pedagogical scaffolds might help teachers effectively use information from 3D assessments (Bybee, 2014). This project will develop PCKSs to help teachers make instructional decisions to effectively promote 3D learning.

| Clar | rify learning Goals | Characterize student performances | Identify academic, social, and cultural background | • | Strategies: Develop Instructional strategies | |
|----------------------------|-------------------------------------|---|--|--------------------|--|----|
| Unp Perf Exp | formance ectation | Develop assessment tasks and diagnostic | Identify target school: Location, class size, accessibility of | Ide | ntify student difficulties and teaching challenge of th group | (|
| SEP | s, and | Identify student | resource | Sea soc | rch learning resources 3D with consideration ial and cultural backgrounds | |
| ider stud | ntify dent | task with considering socio, cultural | student: Ethnicity, prior knowledge, | De din | velop 3D content-specific strategies focusing each nension | |
| with con soci bac | sidering io, cultural kground | Define groups to support 3D; DCIs; | background, learning experiences | Der act stra | velop general strategies and contextualize ivities for implementing the content-specific ategies | |
| Defi kno | ine prior wledge | SEPs and CCCs; Achieved 3D | teacher: prior knowledge, social, cultural background, | Dev cult | velop inclusive strategies for students in a range of tural, social, and academic backgrounds | |
| and | skills | prerequisite knowledge and skills of each | and teaching experiences and orientation | Dev stra | velop instructional strategies (moves) using the ategies | 4 |
| cha | llenge | group | Create teaching scenarios | Fee | edback from internal researchers and expert chers | I |
| ` | * | ``······ | · · · · · · · · · · · · · · · · · · · | | | 5. |

This study was made possible by National Science Foundation grants Award ID: 2101104 (PI Zhai), 2100964 (PI Krajcik), 2101166 (PI Harris), 2101112 (PI Yin). The findings, conclusions, or opinions herein represent the views of the authors and do not necessarily represent the view of personnel affiliated with the National Science Foundation.

nternalizatio

information with

