

Hello, and welcome to video four in a series of videos in which we're describing the scenario-based tasks we developed as part of I-SMART Goal 2, I'm Bob Dolan. Kim Ducharme just gave you a walkthrough in the previous two videos of the co-design process we used with students to develop these testlets and the prototype of that we actually came up with, she would've given you a walkthrough of that. I'll now be talking about the think aloud and interviews that we did with students using these prototypes.

As a reminder, Goal 2 of I-SMART is to design develop and evaluate learning map model-based assessments. In this particular study within Goal 2, we're co-designing and evaluating these testlets for secondary population students, those students without significant cognitive disabilities. We've developed these scenario-based tasks to evaluate a range of depth of knowledge. We're using deeper application to view universal design for learning principles and with a greater emphasis on the formative use of these instructionally embedded testlets.

The research questions that are driving the study is whether students understand how to interact with these new item types, whether they make effective implicit and explicit choices and whether the task seems to assess a range of depth and knowledge. To just go over the task design, again, we have students completing two models. This is the middle school prototype we're describing now. They complete a food web and they complete a causality storyboard. These representing the lower depth of knowledge elements of the task. While they're doing this, they're collecting evidence in this pros and cons clipboard that they're going to be using to build an argument. They'll be using the evidence that they collected together with the models to build an argument in the form of advice tool letter, advice to the governor, a letter of advice to the governor.

The choice that we provided students was in the form of explicit versus implicit. Explicitly in two different ways, first, we gave them a choice of ecosystem to explore and also, we gave them a choice of which evidence they wanted to use in constructing their final argument. The implicit choice was a range of accessibility tools, including read aloud and speech recognition, the use of hints and the use of the glossary. The evidence we collected were utterances and behaviors and interactions during the think aloud portion of the session and then what they said during the semi-structured post-interview.

For analysis, we did an informal structured coding of their utterances in interactions and their responses. We did a cool new take. We did a visual analysis of their interaction patterns, which we'll share with you, and then a basic comparison of their basic and higher order depth of knowledge scores. The students for the study were six middle school students in a rural

Missouri public school, three in grade six, three in grade seven, three male, three female. Three of these students were performing significantly below grade level in science. Three were still performing below grade level, but closer. The sessions were done one-on-one with a researcher, conducted remotely through video conferencing, we used Zoom. We recorded video, audio and screen recording, and the sessions lasted 45 to 60 minutes.

General observations. Well, all the students that did report that they'd covered this content previously within the last year or two, these are COVID times, so of course, it's a little bit hard to figure out what happened exactly when. Students were all highly engaged in the tasks, and on reflection at the end, they believe they did well on the testlets. All but one of the students successfully backed up their argument citing a balance in the ecosystem and cause and effect phenomena in coming up with their final advice. One student was a bit confused, and their argument included both a pro and a con with some confusion.

Also, as a general observation, when we asked students, did you like this test? All responded yes and four of the six used the word fun. We didn't prompt them for this. Some quotes. "Yes, you got to choose your own answer and how you wanted it to go," in reference to the governor's letter. Another student, "Yeah, it went through things step by step, like each question step by step, like the food chain step by step. They liked how we broke things down and scaffolded them in this way." One student responded that there was a bit of a story to it. Again, referring to the writing of the letter. One student specifically said, "Other tests make you read big, long passages," something that we obviously did not include in this task.

We asked students a challenging question. What do you think we would've learned about you based on how you did on this test in terms of what you know, and can do in science? Most of the students really couldn't answer that one. A couple demonstrated some basic metacognitive skill. One said that I'm slow at typing, that I'm relatively smart and that I like rainforest. Another said that I'm not good at spelling, laughed, that I had a few questions where it takes me a while to figure them out.

Let's get into the research questions now. Do students understand how to interact with the new item types? By and large, yes, and quite skillfully. That said, there was a tendency to start their work with regard for instructions, nothing surprising here. They tended to just dive right in and sometimes had to be, especially in the beginning, pulled back a little bit, slow down, what are you seeing? How are you making your decisions about what to do next? Students specifically liked the fact that they were ... that they had to check their answer before more moving on to the next step, and they used that very easily. There was no surprises here to

them. They specifically liked that they had, it gives you a chance to fix your mistakes and you learn from your mistakes. Another student said it's helpful because I understood what I got wrong, which one did I need to work on, and which ones I don't need as much work on.

Again, the immediate response we gave to them was only on the lower-level depth of knowledge type questions. As far as the pros and cons clipboard, they had no problem understanding what it was for, but not exactly why they would use it and they would use it. Also, there was some conceptual confusion just in the nature of the task involving double and triple negatives. When we talk about evidence that there shouldn't be a law protecting jaguars, sometimes these are challenging constructs that students struggled with, but that had nothing to do with the design of the task.

As I mentioned, we came up with an interesting visualization of students' interactions and patterns, behavioral patterns when they went through the task. Here's an overview of it. I'm going to zoom in in a minute, but I just wanted to give you the forest before diving into the trees. Each row is a student. We want to be able to look at the six students as a whole. The columns represent the various screens within the task itself. Here's the welcome screen, here's where they choose an ecosystem, here's where they're reduced to using the clipboard.

These tasks here are where they completed the food webinar, these are where they completed the story board, and then here is at the end where they constructed the argument in the form of a letter to the governor. We also have a screen, it's two columns here in which we've scored their results and a key on the far right, and you'll have access to this as a higher resolution image to play around with yourself. But let me dig in a little bit first to give you some more detail here. We're just going to look at the first student.

These stars indicate where students used some of the built-in tools. Next to the star, it shows which tool they use. Here is a student who used the read aloud, the speech to text, the hints and the glossaries. If you remember in the task, when Kim showed you, we asked them to. We wanted to give them at least a little bit of a chance to try these things out in lieu of proper practice and opportunity to really use these things more regularly. The exclamation points, there's a few of those. Those indicate challenges that the students that the researcher had to intervene for some reason. There was only a few usability blockers. Those are marked here with symbol here. PC indicates that the student used the pros and cons clipboard. Okay.

We've also indicated with colors how the student did on that portion, on that screen. Green means they got everything right, yellow means they didn't get everything right the first time, and the circle around this, the concentric circles indicate how many attempts that they made, and red indicates that they were not able to get this screen correct at all even after multiple attempts. They bottomed out basically. We scored how they did on the lower depth of knowledge tasks on a 10-point scale. We represent those scores here in this column. The evidence that they provided in their advice to the governor was also scored with a rubric we developed on a 10-point scale.

We also have a column indicating what their final screen looked like in terms of the evidence that they provided. All students indicated that we should make a law to protect the species in question, so what we list here is the evidence they provided. Then, on the final columns, are just some of the essentially footnotes here that are indicated elsewhere in the chart. Research question two, did students make effective, explicit and implicit choices? Of course, how do we define effective?

We looked for different types of evidence to evaluate whether we considered that their use of ... whether their making use of these options was effective. In the case of choosing an ecosystem, we asked them afterwards, what did you think about having the ability to choose? Some of their responses were one student said they definitely like being able to choose. "I like projects where you can choose what you want to do and not be forced to do something that you would not want to do." Another student said, "You want to go with the one that feels like it helps you improve." Another student said, "Yeah, being able to choose made a difference, because people might not know about this one ecosystem, but other people might or might not know or not be interested in the other one as much." I think they were referring to themselves here. As an aside, students chose all three of the ecosystems across all six students.

Here, students were able to articulate, from a metacognitive and effective perspective, that they really did like being able to make this choice of ecosystem. As far as being able to choose different types of evidence, two students used three sources of evidence, two students used three sources of evidence, three students used one source of evidence and one student used no evidence. What was interesting to note though is, in the case of one student, they understood the general task well enough that they started sculpting evidence even as they were going through the initial parts of task. For example, during initial use of the pros and cons clipboard, a student thought aloud that coyotes "help regulate the population of whatever they eat." But unfortunately, they didn't record anything of that on the actual clipboard despite looking at the glossary and hints.

Later, they returned to the clipboard, this is after completing the food web, and they were able to articulate that animals, articulate the animals that the coyotes eat and we suspect that when they first articulated this, that they thought that the phrase, "whatever they eat," was too vague, they needed specific examples. They didn't find them until after they completed the food web and then they went and put that into clipboard. But later, they had been integrating this information and forming their final argument. They generalized the food that the coyotes eat to pre and they added that the ecosystem would be less balanced without a law protecting coyotes.

Implicit choice-making. Students generally made little use of hints and glossary, even though they would likely have helped the students. They rarely used the read aloud, even when that would likely help as well. Students often struggled with decoding and comprehending terms such as ecologist, and students were very reticent to use the speech recognition, the speech-to-text, even though when it would've likely helped them. For example, while writing in the clipboard, a student thought aloud a very well-articulated pro argument and we so wished they could have captured that, but instead they really struggled to type in what ended up being much less elaborate response.

What evidence did we collect to evaluate research question 3? Does the task assess a range of depth of knowledge? It's a tough question to ask with such a small sample size, but let's give it a look. Here, we've just plotted students as a single dot. One of these dots corresponds to two students where we're looking at their higher order depth of knowledge scores, a function of their basic depth of knowledge score. We're going to zoom in on three students. Students one and four who scored very well on their basic depth of knowledge elements, but not so well on the higher order depth of knowledge component and student five for whom the opposite was true. If we go back to this pattern, table, students one and four sailed through the food web and storyboard components, each scoring a 9. The final arguments they constructed, eh, not great as you can see.

Contrast that with student five, who had some significant challenges with some aspects of both the food web and the story board, in terms of their ability to demonstrate basic depth of knowledge, but did a phenomenal job at constructing an argument at the end. We've got some evidence here that, yeah, the way we design this does allow us to, it does allow students to independently demonstrate their knowledge and skills tied to specific essential element nodes at the lower DOK level, but for the higher DOK level, they can do that independently. That was just a basic description of some of our findings with a little bit of interpretation, which I'll be saving mostly for the final video what we learned. Thank you.