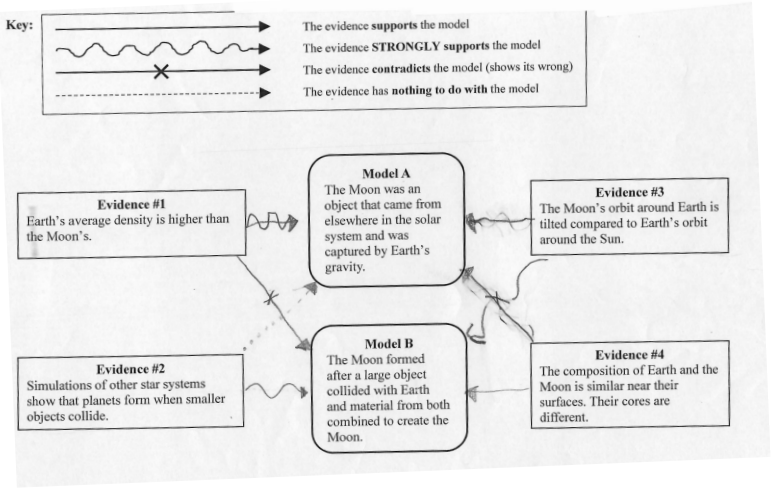
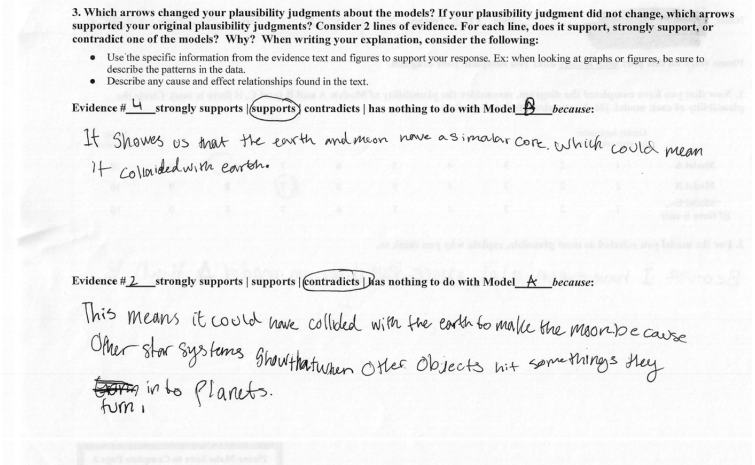
Moon Conspiracy: Relations between the students' written response length and the scientific quality of their responses

Alexandria Wright, Joseph Puig, John R. Robertson, & Doug Lombardi
Human Development Program, University of Maryland

Purpose & Research Context

- MEL diagrams are designed to promote adolescents scientific thinking (Lombardi, et al., 2022)
- Scientific vocabulary may be an indicator of the quality of students' classwork (Roberts & Bybee, 2014; Kuhn, 2011)
    - Scientific vocabulary as an indicator of well-developed scientific literacy (Roberts & Bybee, 2014)
    - Scientific thinking reflects both critical thinking and knowledge of the scientific process (Kuhn, 2011)
- Wanted to see if the number of words in each written responses reflect students' scientific thinking




3. Which arrows changed your plausibility judgments about the models? If your plausibility judgment did not change, which arrows supported your original plausibility judgments? Consider 2 lines of evidence. For each line, does it support, strongly support, or contradict one of the models? Why? When writing your explanation, consider the following:

- Use the specific information from the evidence text and figures to support your response. Ex: when looking at graphs or figures, be sure to describe the patterns in the data.
- Describe any cause and effect relationships found in the text.

Evidence # 4 strongly supports | (supports) | contradicts | has nothing to do with Model B because:

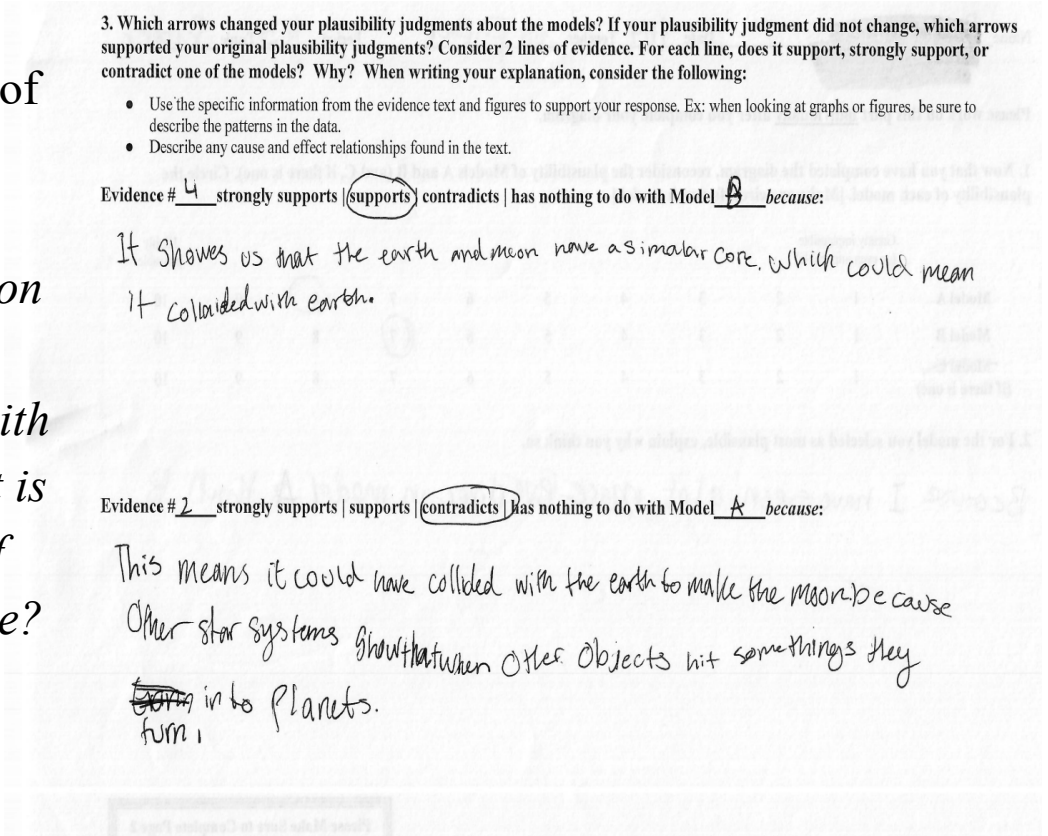
It shows us that the earth and moon have a similar core. Which could mean it collided with earth.

Evidence # 2 strongly supports | supports | (contradicts) | has nothing to do with Model A because:

This means it could have collided with the earth to make the moon because other star systems show that when other objects hit sometimes they form into planets.

Research Question

- We focused on quality and quantity of scientific words and groups of words, which we called quality tokens.
- We wondered: *What is the association between the number of words that students wrote in their explanations with the number of quality tokens and what is the association between the number of words and the quality of their response?*



3. Which arrows changed your plausibility judgments about the models? If your plausibility judgment did not change, which arrows supported your original plausibility judgments? Consider 2 lines of evidence. For each line, does it support, strongly support, or contradict one of the models? Why? When writing your explanation, consider the following:

- Use the specific information from the evidence text and figures to support your response. Ex: when looking at graphs or figures, be sure to describe the patterns in the data.
- Describe any cause and effect relationships found in the text.

Evidence # 4 **strongly supports | (supports) | contradicts | has nothing to do with Model** B ***because:***

It Showes us that the earth and moon have a similar core. which could mean it collided with earth.

Evidence # 2 **strongly supports | supports | (contradicts) | has nothing to do with Model** A ***because:***

This means it could have collided with the earth to make the moon because other star systems show that when other objects hit something they ~~form~~ turn into planets.

Methods

- $N = 24$, grade 6 students from southeast U.S.
- Qualitatively analyzed two written response (A and B), with a prompt asking, “Which judgments changed your plausibility about the explanatory models”
- We each independently counted number of (a) words in the response, (b) “quality tokens,” and (c) quality rating, discussed, and reached consensus (good “counting” reliability, with all ICCs > 0.72)

Student ID	A-3A-Total Words	J-3A-Total Words	C-3A-Total Words	A-3A-Quality Units	J-3A-Quality Units	C-3A-Quality Unit (	A-3A-Quality Rating	J-3A-Quality Rating	C-3A-Quality Rating	A-3A-Comment	J-3A-Comment	A-3B-Total Words	J-3B-Total Words	C-3B-Total Word	A-3B-Quality Wo	J-3B-Quality Wo	C-3B-Quality Uni	A-3B-Quality Rat	3B-Quality Ratin	C-3A-Quality Rating (1 low - 4 high)
640032803	13	13	13	2	3	3	1	3		i do not think the ear	I feel this student, tri	11	12	12	4	4	3	2	2	2
640032813	21	21	21	6	7	6	3	4		makes sense, but i fi	This student used th	21	21	21	6	7	7	3	3	3
640032823	12	12	12	3	4	5	2	2	2	needs more substan	The student just rest	21	21	21	6	6	6	3	3	3
640032833	19	19	19	5	6	5	4	4	3	used accurate evide	This student used th	28	28	28	5	5	5	3	3	3
640032843	26	26	26	3	5	4	1	2		the variables do not	The student wrote a	17	17	17	2	2	2	1	1	1
640032863																				
640032883	15	15	15	4	5	5	4	3		used a good word "c	This student explains	18	18	18	4	4	4	2	2	2
640032893	12	13	13	4	3	5	2	3		doesnt really talk ab	The explanation fron	14	14	14	3	4	4	3	3	3
649932903	36	36	36	7	8	9	4	5		used direct evidence	I think the way this s	37	37	37	5	5	5	2	2	2
640032913	16	16	16	4	4	5	3	2		couldve used more v	This student just give	31	31	31	4	4	4	3	3	3
640032923	45	45	45	3	5	5	3	2		couldve used less w	This student had an	42	42	42	6	6	6	4	4	4
640032933	21	21	21	3	4	6	4	4	4	the student used the	The student had a g	21	21	21	6	6	6	3	3	3
640032943	19	18	19	2	4	7	3	2		could have worded it	This student's explar	5	5	5	2	1	2	2	2	2
640032953	19	19	19	3	3	5	2	2	2	doesnt really talk ab	This student's explar	20	20	20	5	5	5	4	3	4
640032963	22	22	22	4	4	7	2	3		i understand what th	I like how this studer	28	28	28	5	5	5	2	2	2
640032973	40	40	40	4	3	8	2	2	2	doesnt explain why t	The student kind of	22	22	22	4	5	5	2	2	2
640032983	13	13	13	3	4	5	2	2	2	doesnt explain why t	Student's explanatio	14	14	14	3	3	3	2	2	2
640032993	40	40	40	6	7	8	4	4	4	uses evidence from	The explanation was	19	19	19	1	1	1	1	1	1
640033023	16	16	16	3	4	5	4	4	4	used evidence to ex	Student explained w	14	14	14	3	4	5	3	3	3
640033033	37	36	37	3	5	5	4	3		explained that it has	Student uses the evi	28	28	28	4	4	4	4	4	4
640033043	17	17	17	3	3	4	3	3	3	they could have elat	I think the explanatic	14	14	14	3	3	3	2	2	2
640033053	16	16	16	1	2	1	1	1	1	there is no evidence	I do not understand	21	21	21	2	2	2	3	2	3
640033063	17	17	17	4	3	5	2	3		i do not totally under	I think the student ex	22	22	22	5	5	4	4	4	4
640033073	31	31	31	6	6	7	4	5		excellent reasoning.	This student's explar	53	53	53	4	6	12	4	4	4

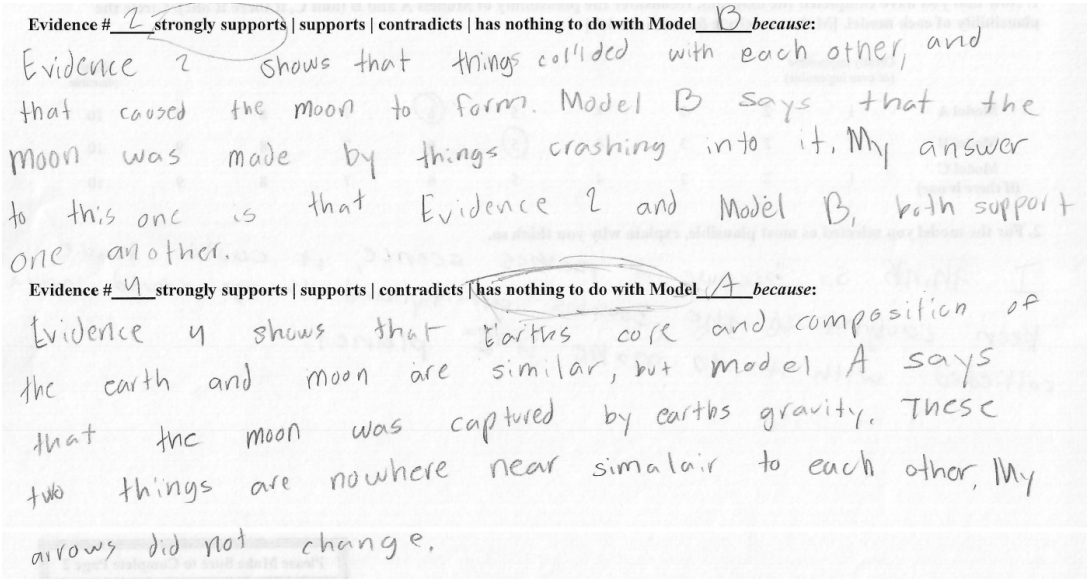
Results

Variable	Total Words 3A	Quality Tokens 3A	Quality Rating 3A	Total Words 3B	Quality Tokens 3B
Total Words- 3A	--				
Quality Tokens- 3A	.543*	--			
Quality Rating- 3A	.256	.506*	--		
Total Words- 3B	.553*	.296	-.161	--	
Quality Tokens- 3B	.167	.297	.079	.714*	--
Quality Rating-3B	.007	-.198	-.274	.472*	.579*

* $p < .05$ (statistically significant correlations)

Discussion

- Higher use of scientific words, associated with high overall quality
- More words could potentially mean more quality tokens
- Future research may include: students' comprehension of the subject when using quality tokens and how much quantity of writing impacts quality



Evidence #2 strongly supports | supports | contradicts | has nothing to do with Model B *because*:

Evidence 2 shows that things collided with each other, and that caused the moon to form. Model B says that the moon was made by things crashing into it. My answer to this one is that Evidence 2 and Model B, both support one another.

Evidence #4 strongly supports | supports | contradicts | has nothing to do with Model A *because*:

Evidence 4 shows that Earths core and composition of the earth and moon are similar, but model A says that the moon was captured by earths gravity. These two things are nowhere near simalair to each other, My arrows did not change.

Alex and Joey's Script For Moon Conspiracy Presentation

Title + Introduction:
Alex:
Hi, my name is Alexandria Wright and I am a third year undergraduate student at the University of Maryland. I am also in the process of obtaining my bachelor of science degree in Human Development.

Joey:
Hello, My name is Joseph Puig. I am a fifth year senior at University of Maryland getting my undergraduate degree in Human Development. I am completing my internship with Doctor Lombardi by working in the SLRG lab with him, along with Alex Wright and John Robertson. The title of our project is, Moon Conspiracy: Relations between the student's written response length and the scientific quality of their responses

Purpose and Research Context:
Our purpose for conducting this research is to examine how early adolescents think scientifically when completing a Model-Evidence Link Activity about how Earth's Moon formed. When examining students' written responses, we wondered "Would adding a necessary set amount of words the students must complete (Min + Max amount of words) in the explanation section help explain their reasoning more?" and if the number of words in each explanation is related to the scientific quality of the explanation. In order to research scientific quality, we had to look at notions from Douglas A. Roberts and Rodger W. Bybee and Deanna Kuhn. In one conceptualization, Roberts and Bybee define scientific literacy as the ability to understand key concepts and principles of science. They call this Vision I, which views science literacy from the outside in, where standards of the scientific process and scientific expertise guide notions of literacy. Kuhn has also explained how scientific thinking arises when a person engages in critical thinking intentionally and goes through scientific processes, such as, data analysis and claim evaluation. In order for something to be described as scientific quality, we conceptualized that the individual looking at these lines of scientific evidence and alternative explanations have engaged in scientific analysis and reflection.

- Vision II: Can be called science for citizenship. Individuals who use this vision believe that science interacts with many human and life endeavors.

Research Question:
We developed our research question after initially reading the 6th graders' explanations. We specifically wondered: **What is the association between the number of words that students use in their explanations with the number of "scientific quality tokens"** (scientifically meaningful words and phrases) **and what is the association between the number of words and the scientific quality of their response?"** Then we conducted an initial round of thematic qualitative coding around the idea of scientific quality in 6th graders' explanations about connections between lines of evidence and models. We examined the quantity, or the number of words in each student's response. We then looked at how many 'scientific quality tokens' the students' responses had and then rated the scientific quality of the students' responses. Overall,

having us look at if the amount of words and quality tokens have an association with the quality of their responses. We were looking to see if more quality tokens would lead to a higher quality response. We chose to focus on the quality and quantity of the student's work because we felt if a student had a greater amount of scientific quality tokens, there may be an associated greater scientific quality in their response.

Methods:
Our sample size for this research was 24, 6th graders from a school in the southeast United States. We made a spreadsheet comparing the first and second explanations (A and B). To begin, we counted the number of total words in each piece of writing. Both Joey and I independently counted the number of words. We then met and discussed any differences in numbers and came to a full consensus agreement. We also independently counted the number of scientific quality tokens, which we described these as "...verbs, adjectives, and nouns used in students' written responses that relate, identify, or make an inference about the meaning of the scientific concept (e.g., the Moon and its formation)". It was difficult to come up with a definition because we would disagree with some of the words that we each considered to be quality. This prompted us to re-examine the literature and come up with a clearer and more precise definition. We independently coded for scientific quality tokens again, compared any differences and came to full agreement for the final count, with good or better inter-rater reliability. After working on the quality tokens, we rated students' written responses on a 1 to 4 Likert-type scale where 1 was low scientific quality and 4 was high scientific quality. Lastly, we made comments about the explanations.

Results:
In the table, 3A and 3B refer to students' first and second written responses, respectively. For response 3A, when the total words for students' response was high, then the quality tokens were also high. We also see that when quality tokens were high, the quality rating was high. These Pearson's bivariate correlations were all greater than .5, which were a moderate effect size. Response 3B was very similar; when the total words for students' response was high, then the quality tokens were also high; and when the quality tokens were high, the quality rating was high. These Pearson's bivariate correlations were all greater than .5, with one above .7, which are moderate to strong effect sizes.

Discussion:
Alex:
After conducting this research, we have much more insight on what goes on in adolescents' brains when thinking scientifically. From the table, higher use of scientific words was associated with high overall quality. This may mean that teachers should encourage students to use more quality words when reflecting and writing. This could potentially help them gain more understanding of complex topics, such as how the Earth's Moon formed. This then raises the question of do students actually comprehend what they are writing about when using quality tokens? Future research can build on a question like this.

Joey:
The research also shows that when students usually write more, their word quality is greater. Our research also suggests that a student's writing quality may be better when they use more quality words. This could imply that when teachers are teaching they should encourage their students to write and elaborate more on their response and explanations because it will help increase their quality of writing and their experience. Teachers may also wish to read their student's work more thoroughly to look for scientific quality words in their responses. More future research should be done though in order to get more information on how effective the **quantity** of writing can be on a student's **quality** of writing.

Token: word or combination of words that are semantically meaningful
Quality tokens def: vocab from the readings we had