EXAMINING THE DEVELOPMENT OF FIFTH AND SIXTH GRADE STUDENTS' EPISTEMIC CONSIDERATIONS OVER TIME THROUGH AN AUTOMATED ANALYSIS OF EMBEDDED ASSESSMENTS

Joshua M. Rosenberg and Christina V. Schwarz
Michigan State University

April 14, 2016
National Association for Research in Science Teaching Annual International Conference

Background

- Focus on learning about scientific concepts through engaging in scientific and engineering practices
- Developing epistemic considerations in classroom settings over a long period of time may be challenging for teachers and learners

(Berland, Schwarz, Krist, Kenyon, Lo, & Reiser, advance online publication; National Research Council, 2012; NGSS Lead States, 2013)

Background

- Understanding children's epistemic considerations can be challenging
  - Contextualized (in *practice*)
  - May take awhile to develop
  - Coding can be labor-intensive

Background

- Automated approaches to analyzing text data have increasingly been used in science education
  - ossible to examine conceptual aspects of students' transcribed responses
  - Embedded assessments may be amenable to text analysis
  - Exploratory approach can examine knowledge *in situ*
- However, researchers have not yet examined epistemic considerations longitudinally

(Beggrow, Ha, Nehm, Pearl, & Boone, 2014; Sherin, 2013; Guo, Xing, & Lee, 2015)

Background

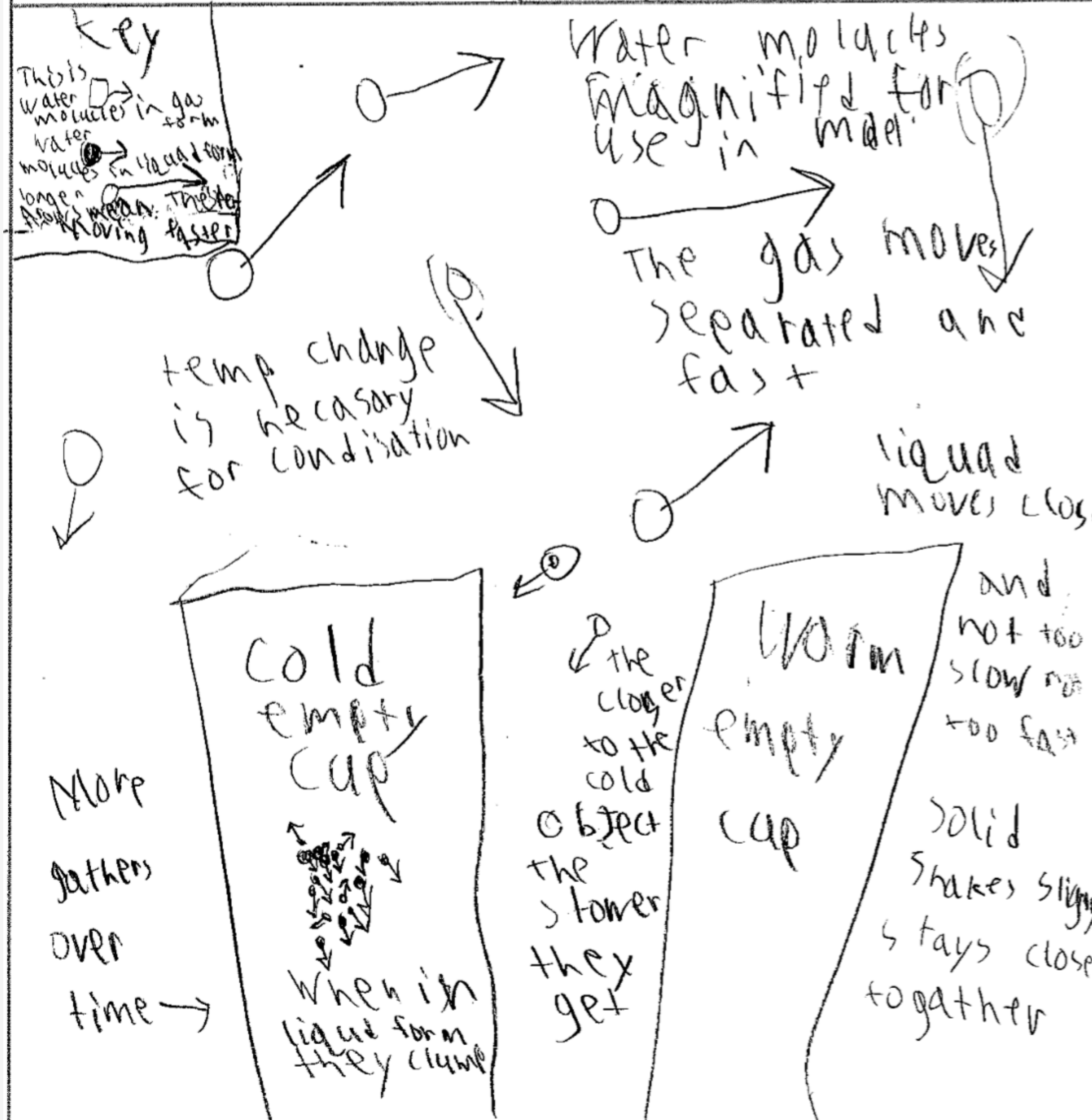
- **Purpose**: Understand what themes can be identified in students' epistemic considerations through analyzing embedded assessments
- If meaningful, examine patterns of themes over time

Method

- Utilized responses from a subset (43) of students taught by one of two fifth-grade and two-sixth grade teachers
- Collected 200 embedded assessments from six units
  - Each included a prompt
  - Each included eight-10 items on epistemic considerations and “meta” items about scientific practices
  - Analyzed six items consistent across all six units

2nd Model of Condensation

Name (initials)		Date (MM/DD/YY)	



Method

- Epistemic considerations
  - Nature
  - Audience of model
  - Justification
  - Generality
  - (Meta / reflective)

Method

- Audience of model
  - Who do you think your model is for?
- Generality
  - Do you think your model should explain all the different ways that [specific to unit] or should it mainly focus on a specific situation like [specific to unit]?

Method

- 5th Grade Units
  - Evaporation (~1 month)
  - Condensation (~1 month)
  - Light (~3 months)
- 6th Grade Units
  - Chemistry I (~1.5 months)
  - Chemistry II (~1.5 months)
  - Earth Science (~2 months)

Method

- 5th Grade Units
  - Evaporation (~1 month)
  - Condensation (~1 month)
    - Draw or attach a copy of your revised condensation model to answer the question: “How and why do liquids sometimes appear on cold surface over time?”
  - Light (~3 months)
- 6th Grade Units
  - Chemistry I (~1.5 months)
  - Chemistry II (~1.5 months)
    - Draw or attach a copy of your individual revised model that answers the question: “How and why do odors move across the room?”
  - Earth Science (~2 months)

Method

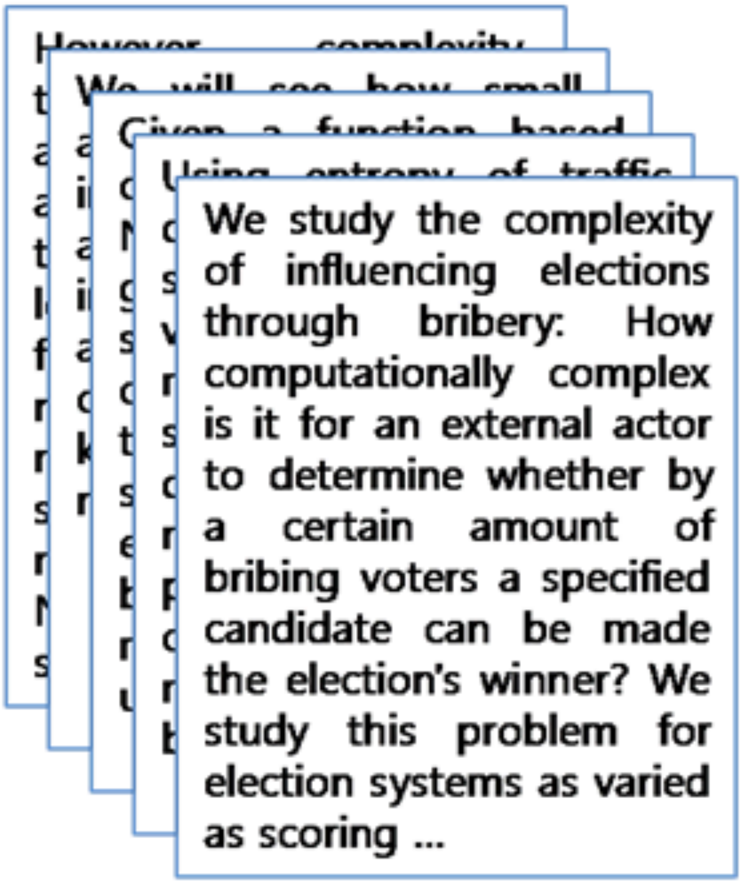
- Adapted Statistical Natural Language Processing technique described by Sherin (2013)
  - Focus on epistemic aspects
  - Analysis of a moderately-sized sample instead of individual students
  - Length of responses

Method

- Entered responses to six items for 200 embedded assessments from 43 students
- Cleaned text and removed a small number of *stopwords* using the tm package in the statistical software and programming language R
- Created term document matrices or vector-space representation

(Feinerer & Hornik, 2015; R Core Development Team, 2016)

Documents



Vector-space representation

	D1	D2	D3	D4	D5
complexity	2		3	2	3
algorithm	3			4	4
entropy	1			2	
traffic		2	3		
network		1	4		

Term-document matrix

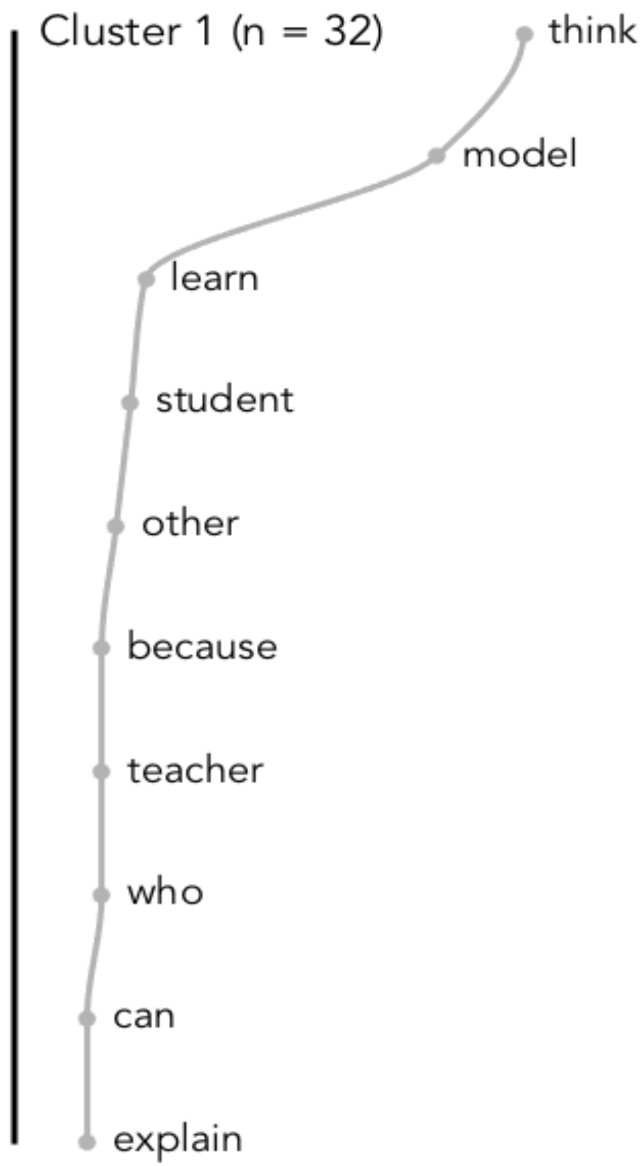
Choi (2016)

Method

- Selected the number of clusters (or themes)
- Clustered documents using a two-step approach
  - Hierarchical
  - K-means
- Interpreted clusters inductively from the data
  - Inspected mean term frequencies and documents for each cluster
  - Examined frequencies of clusters over time

Audience

I think my model is for other students

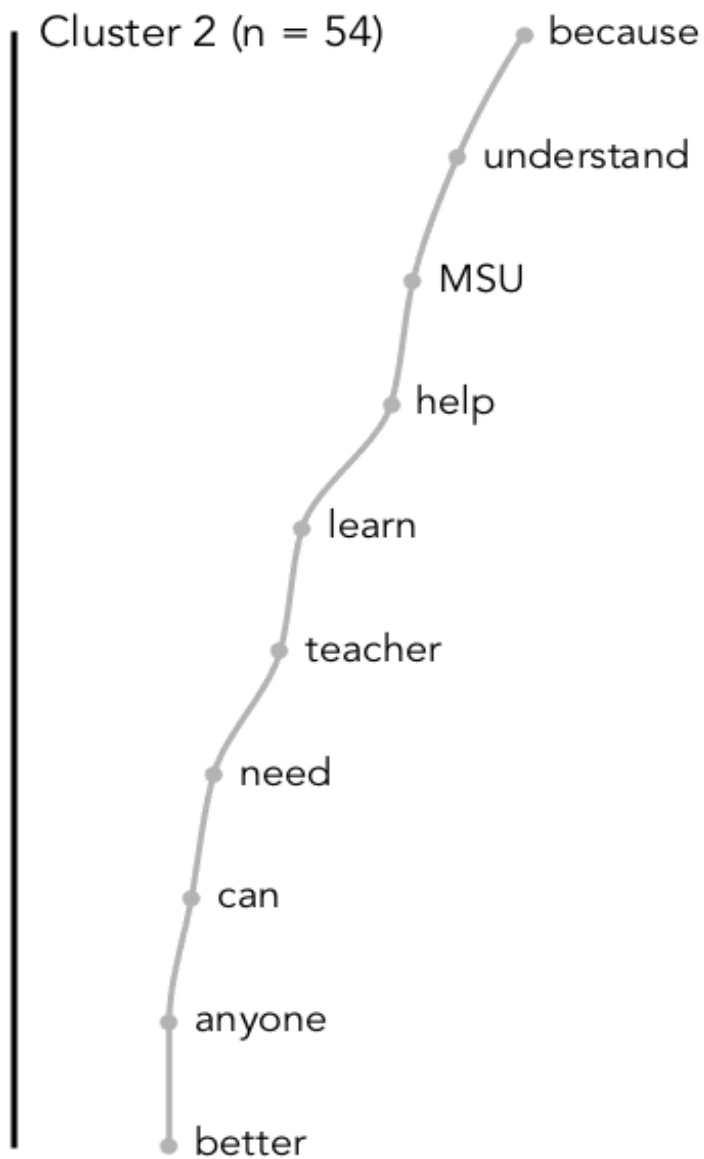


"I think my model is for my teacher and other students."

"I think it is for other students. This is because it helps other students learn about how water shapes our world or we could compare what we think."

"I think my model is for my class and my self."

MSU

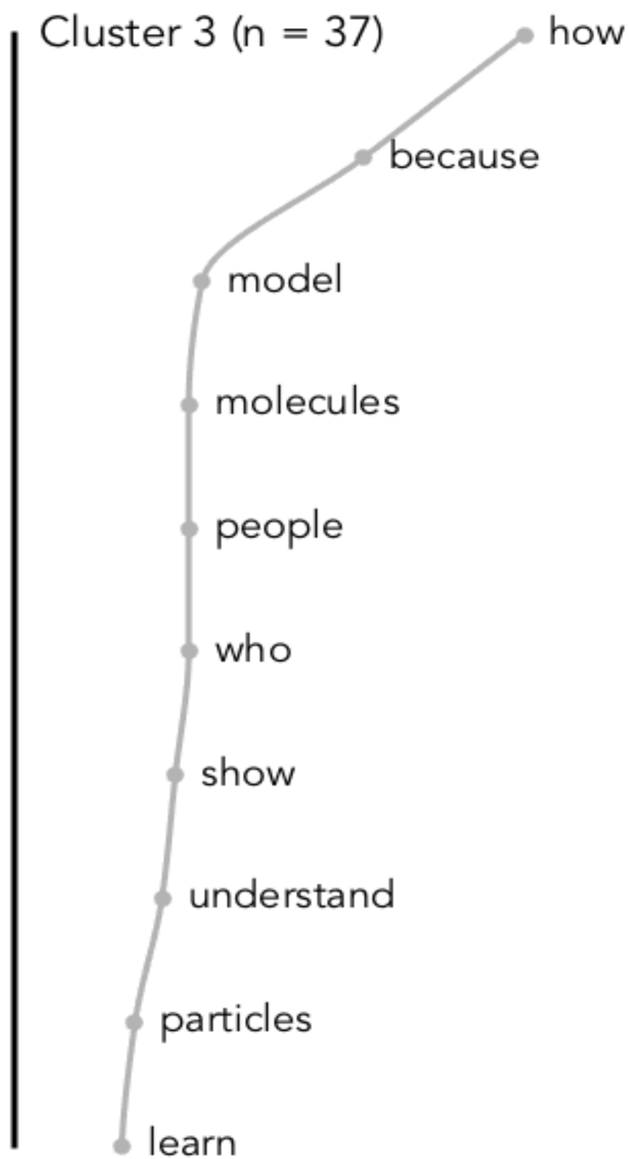


"For the MSU research group and myself its for myself so I can understand condensation better."

"Me and MSU. To teach me and for MSU to research."

"To learn from and help me understand. Because it helps me understand better when we do more and more."

For people to understand how

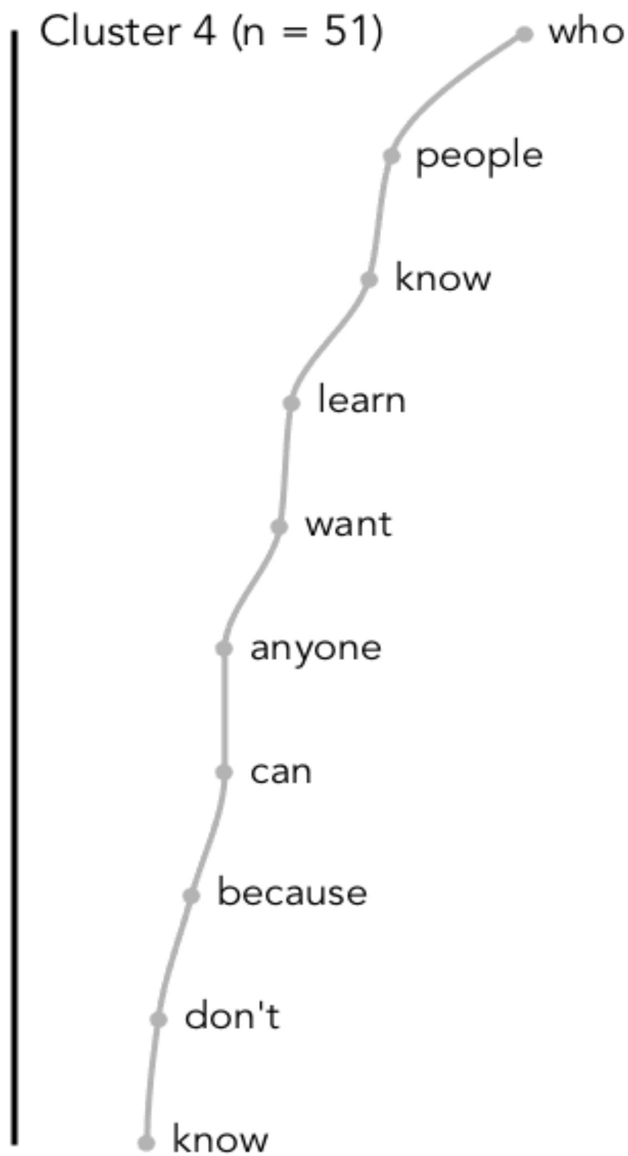


"People who don't understand because then they can look at my model and see how it works."

"For people who want to know how you see [some]thing. Because that's what the model is for."

"People who don't understand ideas about odors, molecules, and movement. Because then they will partly understand how odors move and what happens to odors."

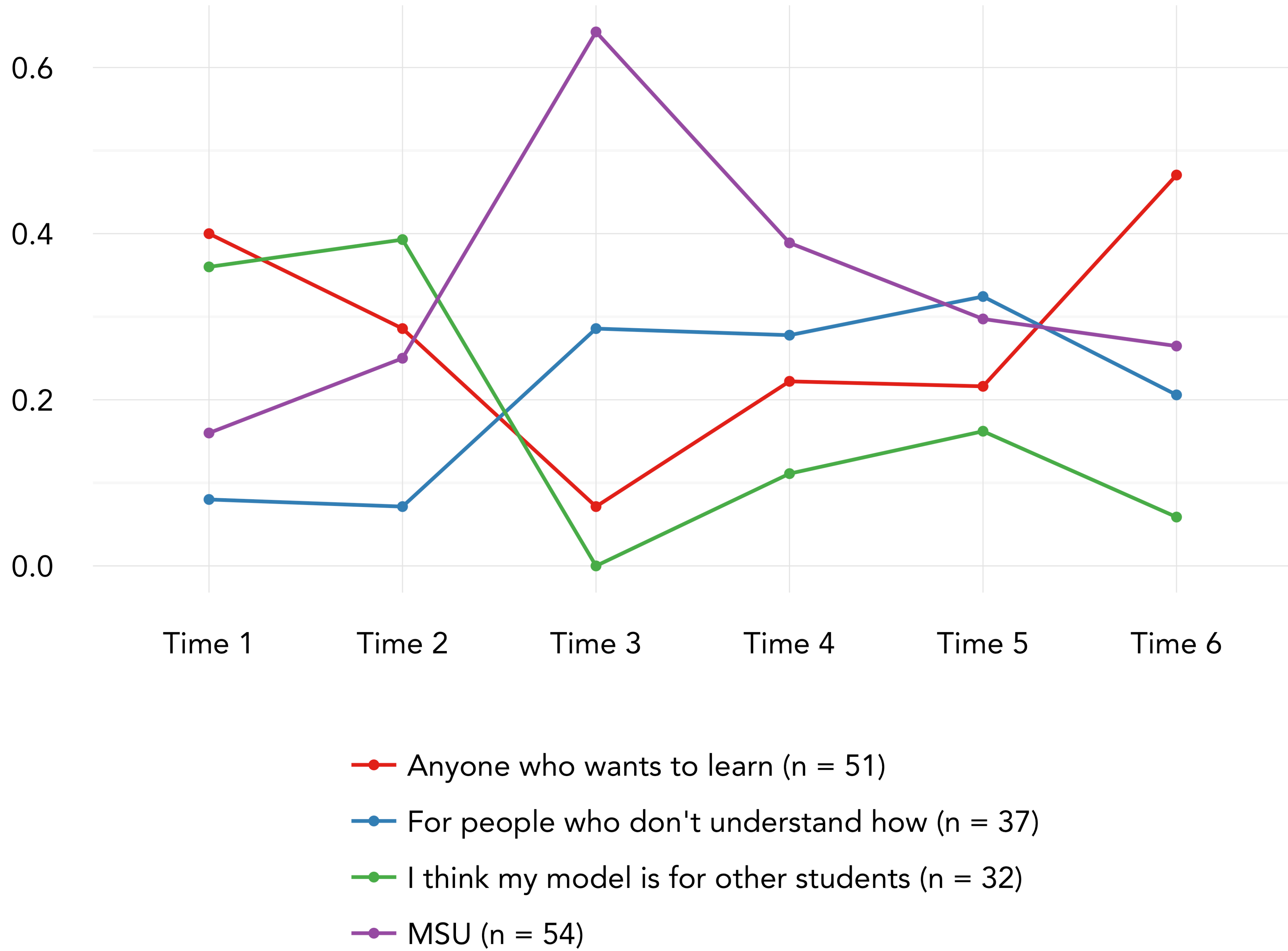
For anyone who wants to learn

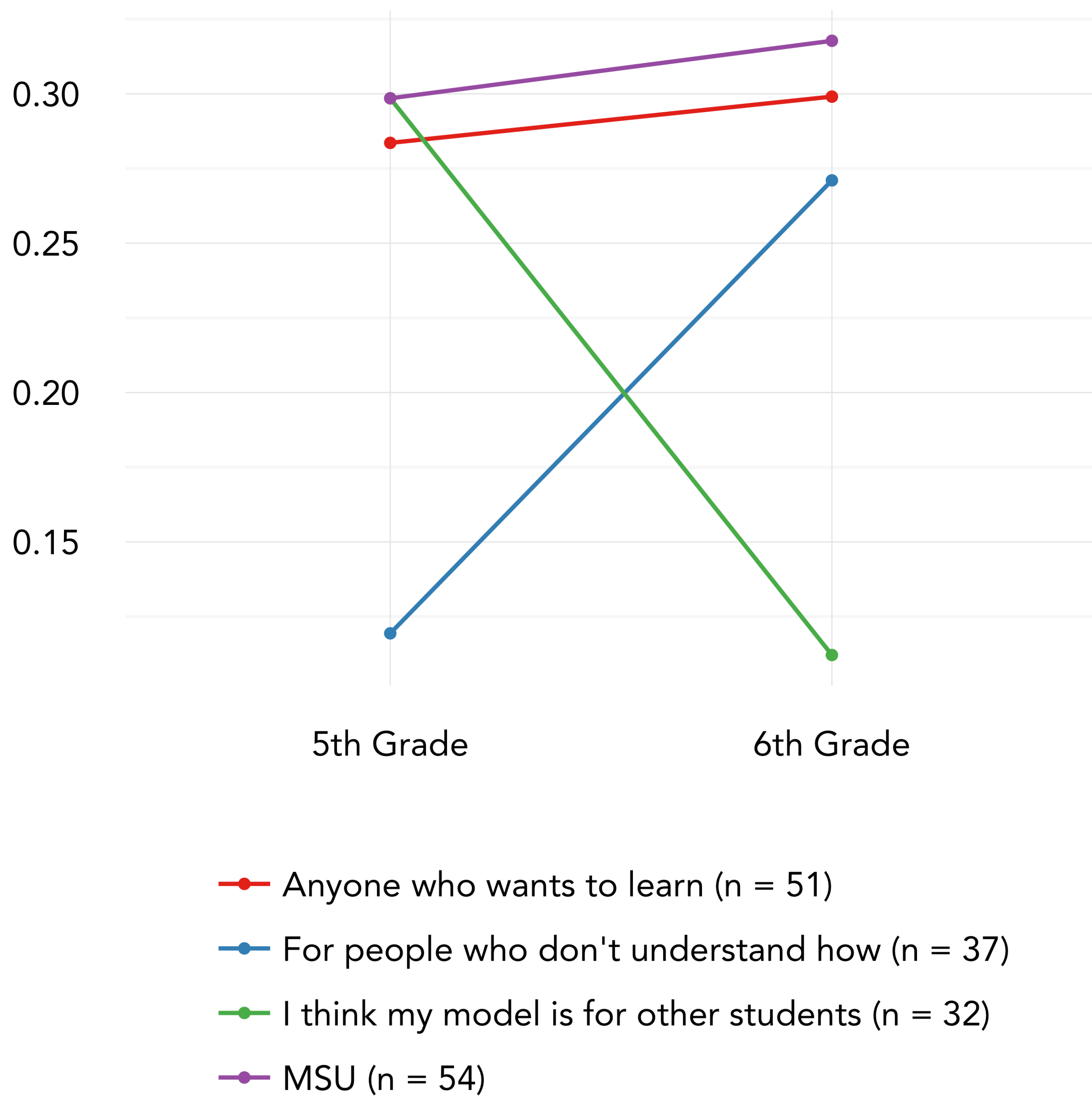


"Anyone who wants to learn about condensation."

"It is for anyone who wants to learn about this kind of stuff. Because people could look at my model and learn about air molecules."

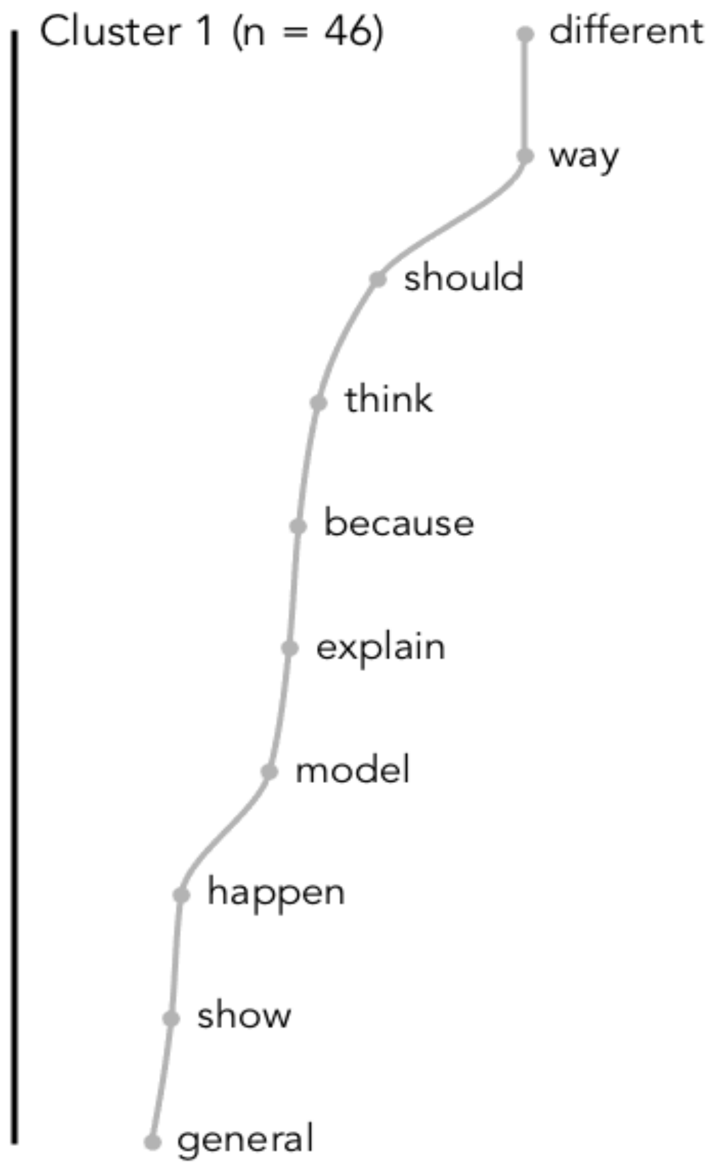
"Anyone who wants/needs to know about odor. Because it is an informative model to inform people."


0.6
0.4
0.2
0.0
Time 1
Time 2
Time 3
Time 4
Time 5
Time 6
Anyone who wants to learn (n = 51)
For people who don't understand how (n = 37)
I think my model is for other students (n = 32)
MSU (n = 54)


0.30
0.25
0.20
0.15
5th Grade
6th Grade
Anyone who wants to learn (n = 51)
For people who don't understand how (n = 37)
I think my model is for other students (n = 32)
MSU (n = 54)

Generality

Explain all different ways

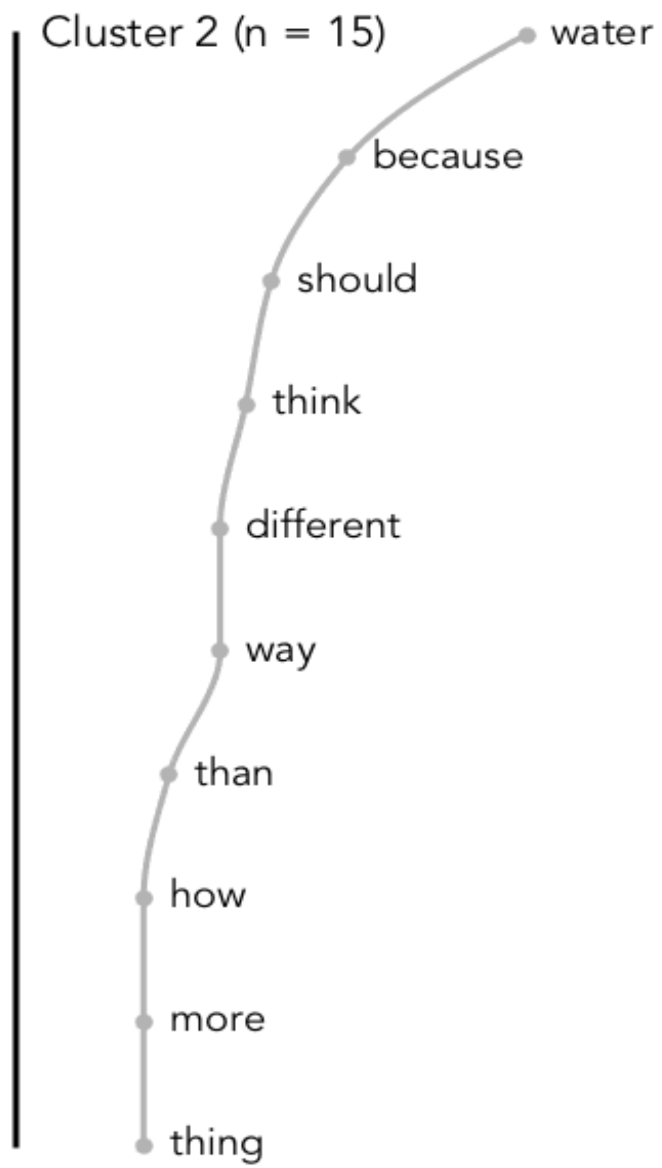


"All different ways. Because that is not the only way evaporation happens. A little child might think it is if it focuses only on one phenomena."

"I think it should explain different ways that evaporation happens. Because it has to explain evaporation, the big idea, and has to show all the kinds of evaporation."

"All the different ways. A good model is general."

Ways water shapes things

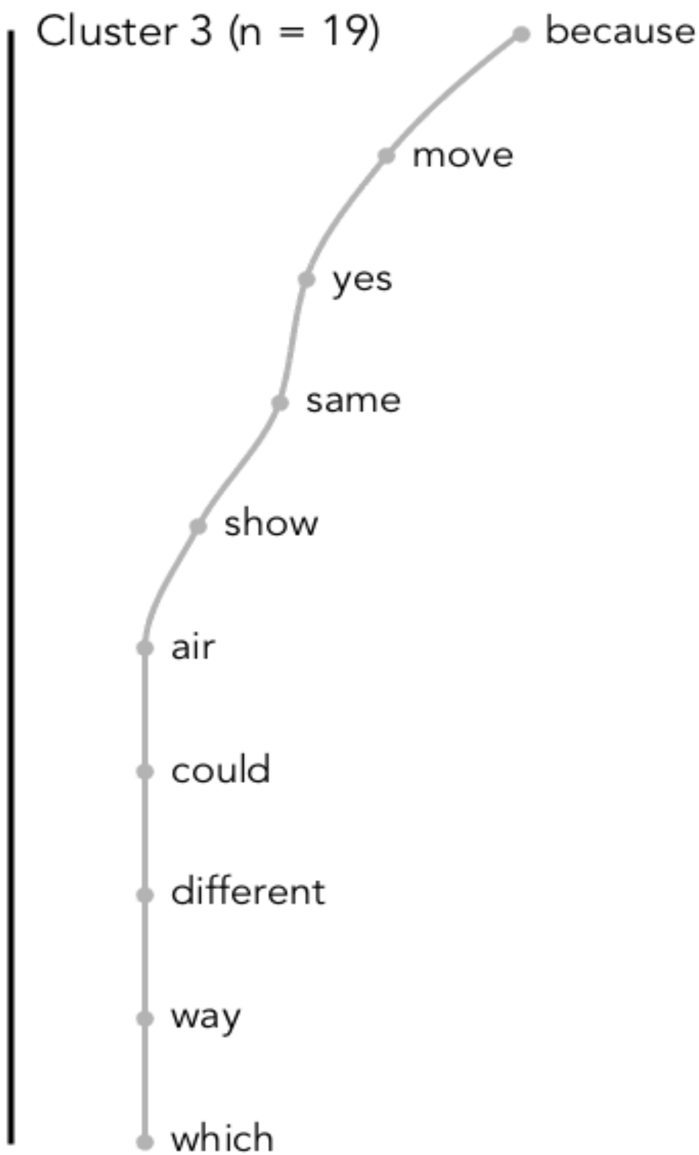


"I think it should show all the way water shapes things."

"All things because the Grand Canyon isn't the only thing that water formed."

"Because water forms more than one thing. Because the water explains how some landforms are formed."

Show the way air moves

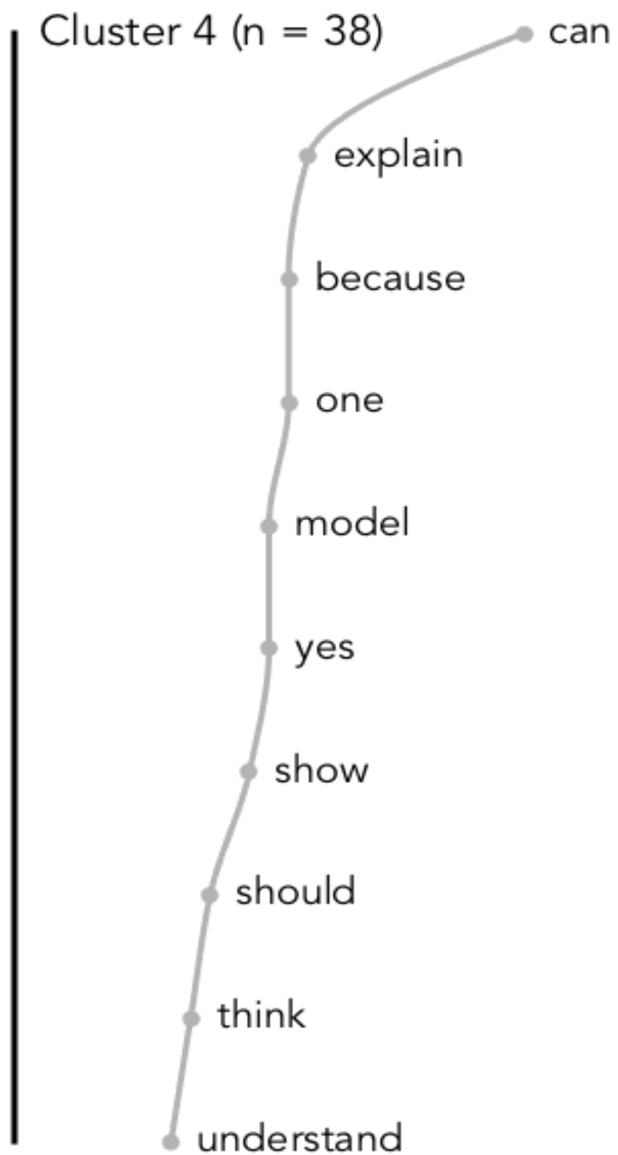


"My model works for all molecules in general. All air molecules and odor vapors move the same. The difference would be seen if you drew specific molecules."

"Yes, because the air molecules could represent any smell. It could be perfume, air fresher, etc."

"Yes because all odors move the same."

Can explain one thing

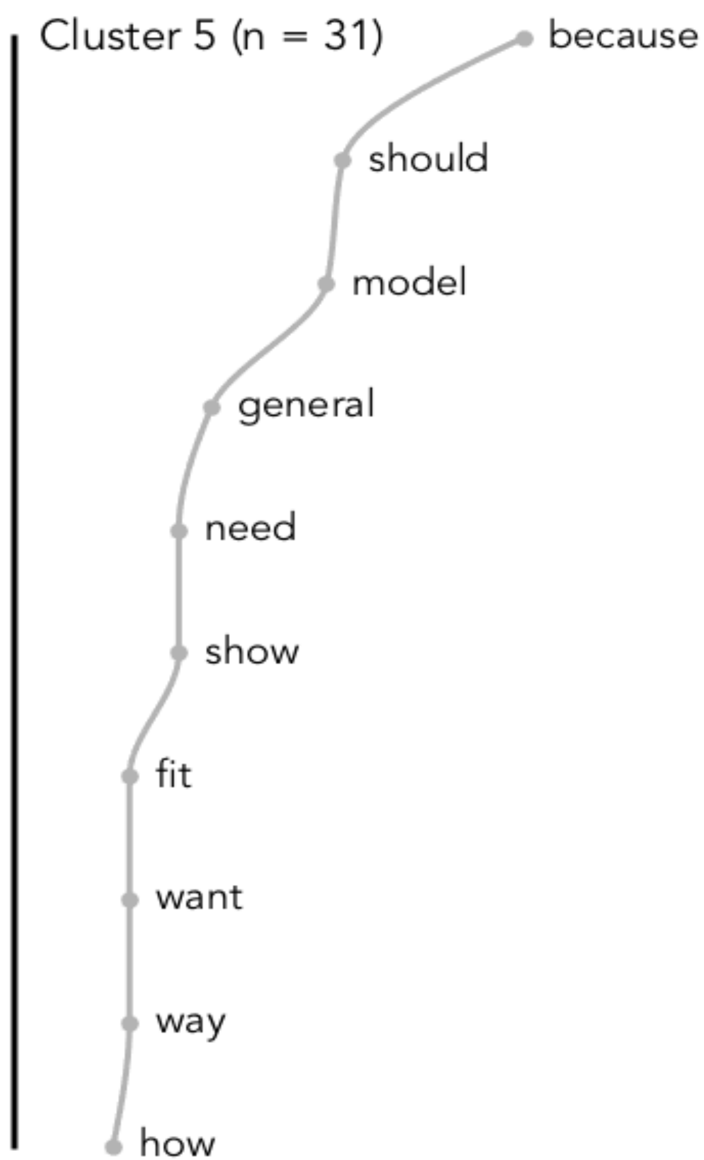


"It should teach on one thing. It is easier to explain and that you can put one thing in more detail."

"Only the cold pop can and ice pack because they shouldn't see every thing in one model."

"It should explain all the types of evaporation. Because then it would be better instead of showing so many models you can just show one."

Models should be general

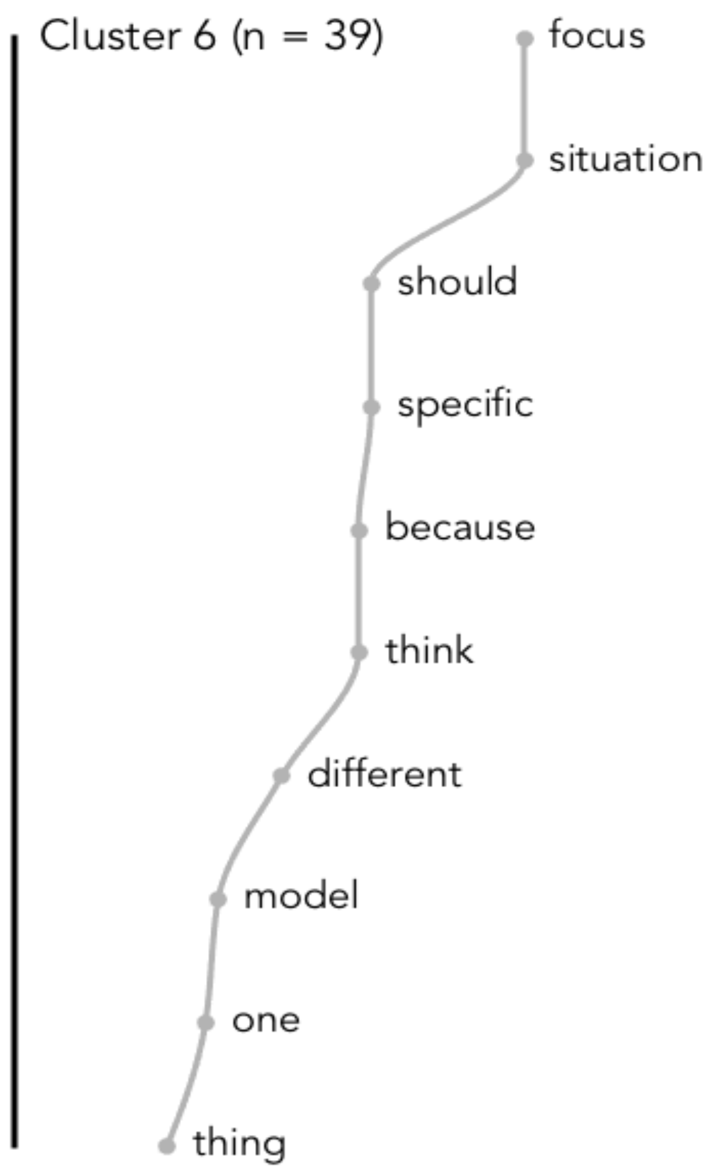


"Not be so specific. Because all good models should fit all phenomena."

"Not too specific. A good model is general."

"My model should explain something in the middle. My model should explain something in the middle because a model should be general, but not so general that it becomes inaccurate for describing some phenomena."

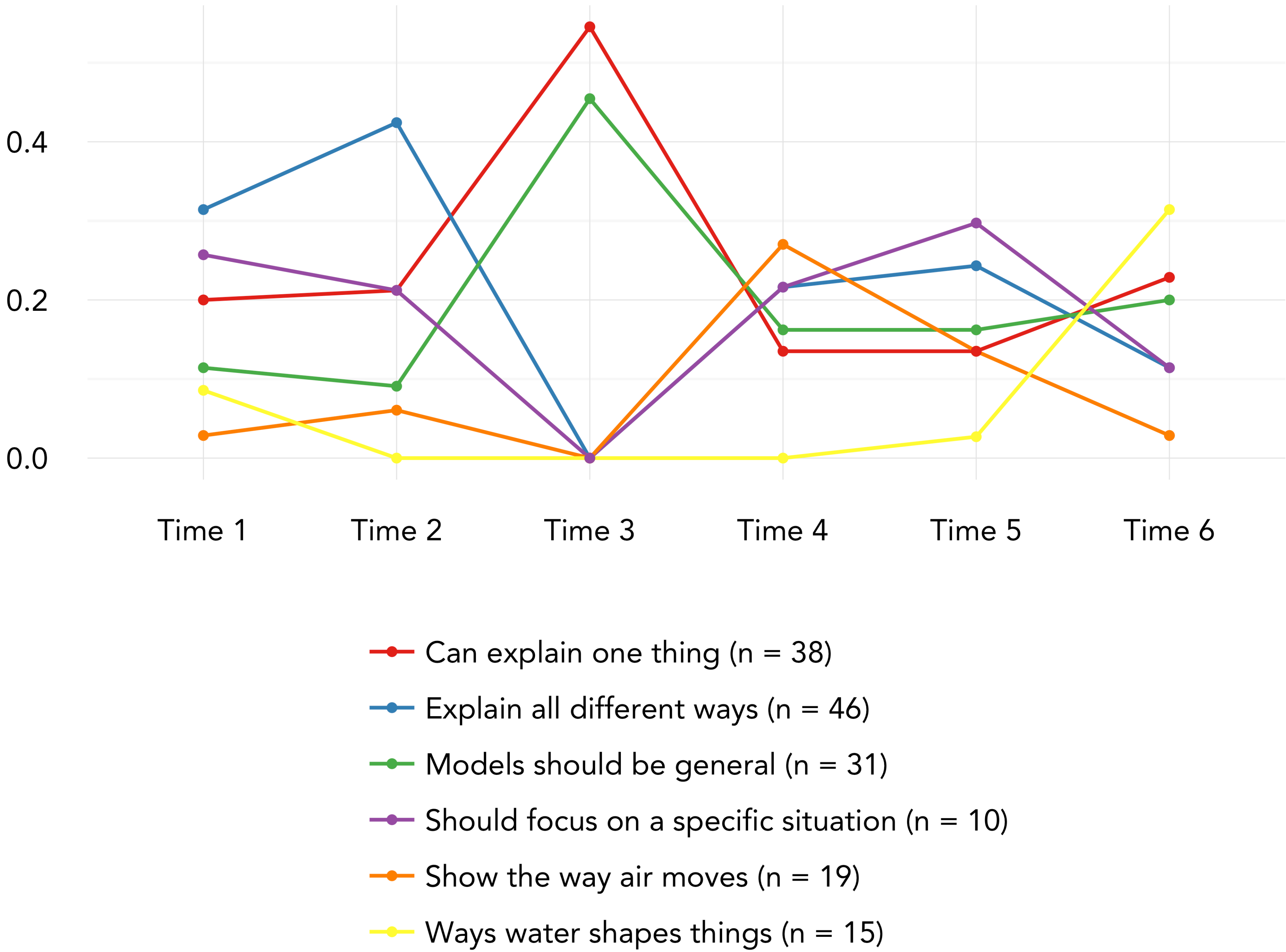
Should focus on a specific situation

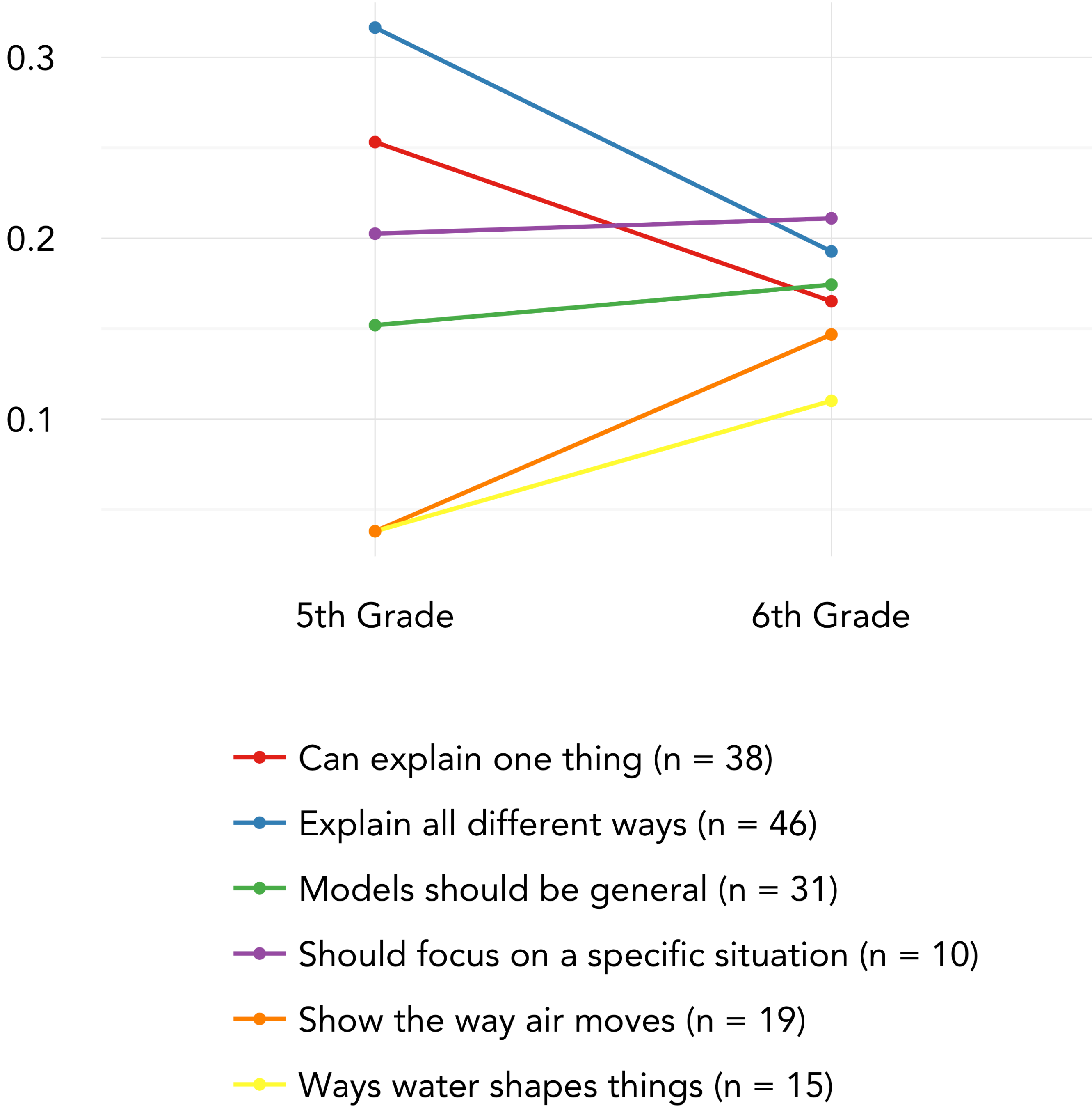


"I think it should focus on a specific situation. If you focus on multiple things it will look messy and it will be hard to read."

"I think my model should mainly focus on a specific situation. Because then it doesn't go off in a bunch of different directions and get confusing."

"I think the model should focus on the big idea (evaporation). Because if you describe to much of one thing you start going away from the big idea."


0.4
0.2
0.0
Time 1
Time 2
Time 3
Time 4
Time 5
Time 6
Can explain one thing (n = 38)
Explain all different ways (n = 46)
Models should be general (n = 31)
Should focus on a specific situation (n = 10)
Show the way air moves (n = 19)
Ways water shapes things (n = 15)


0.3
0.2
0.1
5th Grade
6th Grade
Can explain one thing (n = 38)
Explain all different ways (n = 46)
Models should be general (n = 31)
Should focus on a specific situation (n = 10)
Show the way air moves (n = 19)
Ways water shapes things (n = 15)

Key Findings

- Themes from the automated analysis seem to pick up on different dimensions
- Audience
  - Seems to be highly interpretable but procedural
- Generality
  - Seems to be content-specific of focused on being either general or specific

Key Findings

- Longitudinal patterns demonstrate trends in themes that might be meaningful
- Audience
  - Some growth over time
- Generality
  - More challenging to interpret

Significance and Limitations

- Yes (students) can!
  - Students are responding with not only their epistemic considerations but also others
- We can, too
  - Suggests epistemic considerations and patterns over time can be examined
- But, significant methodological challenges
  - Significant variability within clusters
  - Importance of factors in addition to time
  - Need for validation

Future Directions

- Code additional embedded assessment responses
- Include other data sources to substantiate findings or to serve as factors in addition to time
- Combine classification with clustering
- Focusing on stopword removal to focus on epistemic (rather than procedural or content) aspects

Thank You and Contact Information

- Collaborating teachers and students, the Scientific Practices Research Group, and the National Science Foundation (DRL 1020316)

- Contact:
  - Joshua Rosenberg
    - jrosen@msu.edu
    - http://jmichaelrosenberg.com
  - Christina V. Schwarz
    - cschwarz@msu.edu
    - http://schwarz.wiki.educ.msu.edu/

References

Beggrow, E. P., Ha, M., Nehm, R. H., Pearl, D., & Boone, W. J. (2014). Assessing scientific practices using machine-learning methods: How closely do they match clinical interview performance? Journal of Science Education and Technology, 23(1), 160-182.

Chinn, C. A., Buckland, L. A., & Samarapungavan, A. L. A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. Educational Psychologist, 46(3), 141-167

Ingo Feinerer and Kurt Hornik (2015). tm: Text Mining Package. R package version 0.6-2. http://CRAN.R-project.org/package=tm

Guo, Y., Xing, W., & Lee, H. S. (2016). Identifying Students' Mechanistic Explanations in Textual Responses to Science Questions with Association Rule Mining. 2015 IEEE International Conference, Atlantic City, NJ. 10.1109/ICDMW.2015.225

R Core Team (2015).

Sherin, B. (2013). A computational study of commonsense science: An exploration in the automated analysis of clinical interview data. Journal of the Learning Sciences, 22, 600-638.

R Development Core Team (2016). R: A language and environment for

statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.