

Clustering Analysis Reveals Authentic Science Inquiry Trajectories Among Undergraduates

Melanie Pepper

University of Northern Colorado
Ross Hall 1556 Campus Box 92
Greeley, Colorado, 80639 USA
melanie.pepper@unco.edu

David Quigley

University of Colorado Boulder
1777 Exposition Drive
Boulder, Colorado, 80301 USA
david.quigley@colorado.edu

Mehrgan Mostowfi

University of Northern Colorado
Kepner Hall 0095C Campus Box 128
Greeley, Colorado, 80639 USA
mehrgan.mostowfi@unco.edu

ABSTRACT

Science education reforms in the United States call for an emphasis on teaching of scientific practices, such as inquiry. Previous work examined expert versus novice practices in authentic science inquiry and found although experts have fairly consistent inquiry strategies, novices exist on a continuum. In this paper, we extend our previous qualitative work to quantitatively analyze differences in inquiry practices among novices. Using clustering analysis, we found that non-science majors who performed simple investigations tended to cluster together and biology majors who performed complex investigations also tended to cluster together. We observed two additional clusters that contain both non-science majors and biology majors, but who performed distinct inquiry strategies. This raises some critical questions about how to pedagogically target students within each cluster.

CCS CONCEPTS

• Applied computing~Education • Applied computing~Law, social and behavioral sciences~psychology

KEYWORDS

Authentic science inquiry, science practices, science classroom inquiry simulations, clustering

ACM Reference Format:

Pepper, M., Quigley, D., and Mostowfi, M. 2019. Clustering Analysis Reveals Authentic Science Inquiry Trajectories Among Undergraduates. In *The 9th International Learning Analytics and Knowledge Conference (LAK19)*, March, 2019, Tempe, AZ, USA. ACM, New York, USA. 5 pages. <https://doi.org/10.1145/3303772.3303831>

1 INTRODUCTION

1.1 Science Inquiry as a Practice

Recent science education reforms in the United States push for teaching students not only disciplinary knowledge, but a knowledge of science practices [13]. Scientific practices include competencies such as argumentation, model building,

and inquiry. Inquiry can be classified as either simple or authentic in nature, with simple inquiry prevailing in K-16 settings [4, 15]. Although simple inquiry is more accessible in the classroom, previous work has suggested that engagement in only simple inquiry can negatively impact students' science literacy, as students believe that science is a straightforward, easy process that exists only to generate facts [4].

In response to these concerns, new technologies and pedagogical approaches have been developed to provide more scientifically authentic inquiry experiences to students. For example, Course-Based Undergraduate Research Experiences or CUREs are gaining in popularity as a method for giving undergraduate students authentic science experiences as part of their required lab coursework [3, 6]. Science Classroom Inquiry (SCI) simulations place the student in the role of a scientist tasked with uncovering the cause of some real-world phenomena or problem. The authenticity of the experience is derived not only from the use of real-world data, but the lack of a solution for the students to uncover, the ability to revise their hypothesis, freedom to pursue any testing strategy, and the ability to conclude when the students feels that they have collected enough data. Previous work by [17] demonstrated that in SCI, students perform investigations reminiscent of simple inquiry, as defined by [4], or highly complicated investigations that are reflective of what practicing scientists do in real life and in this simulated environment. Why do some students perform more scientifically authentic investigations than others? Differences in practices may reflect differences in how students understand the nature and process of science and science knowledge, or their epistemological beliefs about science.

1.2 Inquiry Practices and Epistemology

Alongside practices and disciplinary knowledge, developing sophisticated epistemological beliefs about science is a cornerstone of science literacy [20]. Although epistemological beliefs about science, and the related construct, nature of science (NOS) understanding, are both recognized as important for science literacy, they are notoriously difficult to define and assess. For example, practicing scientists are not consistent with how they operationalize NOS [22, 24], there is a lack of consensus concerning if NOS is universal or varies depending on discipline [1, 23], and there is a lack of

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

LAK19, March, 2019, Tempe, AZ, USA
© 2019 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-6256-6/19/03.
<https://doi.org/10.1145/3303772.3303831>

agreement on what it means for an epistemological belief to be “sophisticated.” With regards to assessment, using standard survey metrics, we further constrain our understanding of this important phenomena by examining what an individual knows or believes at only one point in time. An emerging solution is to examine student practices in real time and authentic disciplinary contexts, such as argumentation [8] or inquiry [17, 21]. Previous work by [17] extends this argument to suggest that student inquiry practices may also be reflective of underlying epistemological beliefs and/or NOS understanding which was termed Epistemology in Authentic Science Inquiry, or EASI. EASI describes an individual’s NOS understanding and epistemological beliefs about science as reflected by inquiry practices situated in the SCI environment. Student inquiry practices can be captured and analyzed using learning analytics techniques. For example, [16] used natural language processing tools to analyze the language used by experts or novices (where expertise was defined by the user’s experience with authentic science practices) during the conclusion phases of a SCI simulation. The authors noted that experts used tentative language when making their conclusions whereas novices used more definitive language [16]. Since understanding that scientific knowledge is tenuous and subject to revision in response to new evidence is a sophisticated epistemological belief, this study supports the proposition that learning analytics tools can be utilized in authentic settings to assess amorphous constructs such as epistemology. In this study, we expanded on our previous qualitative work by using clustering techniques to capture patterns found in users’ clickstreams as a high-throughput method for analyzing differences in inquiry approach.

2 METHODOLOGY

2.1 Participants

106 undergraduate students were recruited from a mid-sized, public, doctorate-granting institution located in a small city located in the rocky mountain region of the United States. The participant population was predominantly female (71.3%). 56.52% were non-STEM majors hailing from programs such as psychology, business, and criminal justice and 43.48% were biology majors.

2.2 Data collection

All data was collected during a single meeting. Participants were logged into the *Invasion of the Grackles* SCI simulation. In this simulation, students were tasked with determining a cause for the range expansion of a nuisance bird species, the Great-tailed Grackle into Northern Colorado. Students were given complete autonomy as to the generation of their hypothesis (H), how many and which tests (T) to perform in the order of their choosing, and whether or not to seek outside information (I) either through the in-simulation library or via the Internet. Participants had autonomy to

decide when they were finished and generate their final conclusion (C). Actions made in the simulation were captured by the SCI web application [15] for analysis and actions outside the simulation were captured by Windows Steps Recorder. Demographic information was captured after the user completed their investigation to avoid stereotype threat.

2.3 Data analysis

2.3.1 Scoring. As another measure of scientific practices, we also examined the log files generated by each person as they completed the simulations. These simulation “notebooks” include not only the clickstream data, but the user’s written rationale for each step, hypotheses generation, and final conclusions. Based on the coding strategy in [17] each notebook was coded as simple or complex. Simple users were not systematic or mechanistic in their approach and their overall inquiry strategy was reminiscent of simple inquiry as described by [4]. Complex investigations were logical in nature, evidence-based, and final conclusions were often mechanistic in nature. Notebooks were coded blind by two individuals and disagreements settled through discussion. Cronbach’s alpha was 0.71, indicating acceptable agreement.

2.3.2 Clustering. Similar to [19], we normalized each user’s clickstream to the series of investigation practices they used within the SCI tool. We broke the clickstreams into relevant features. Data was then merged and normalized as investigation actions (H, T, I, C) to generate a string of actions performed by each user. Once an action sequence was generated for each user, we transformed this stream into a series of 14 features to capture patterns within the data. These features parallel the qualitative differences among experts and novices as described in [17]. The exact details of these features are outlined in Table 1.

Table 1: Feature Set Selection

Feature Set	Features	Mapping
Raw Action Count	Number of H, I, T, C; Actions before initial H	In prior work [6, 18] experts tend to do more information seeking actions than novices.
Iteration Features	Max length of repeated I or T; Number of action transitions after initial H	In prior work [6] experts tend to perform information seeking interspersed with testing. Novices will be more likely to follow a single line of inquiry without seeking additional information.
Frequencies	Relative frequency of H, I, T, C	In prior work [6], experts tend to perform more tests and information seeking.

Once we developed a feature representation of each user’s actions, we used Weka [10] to generate a series of user clusters based on different feature subsets using K-Means. We used the elbow method [12] to determine the appropriate number of clusters in each analysis. By the elbow method, the within-cluster variance (coordinate-wise squared deviations

from the mean of the cluster of all the observations belonging to that cluster) and the determination that the number of clusters is appropriate is made when the variance does not change significantly by adding another cluster. In our cases, the variance “elbowed” when the number of clusters was four. We examined the distribution of centroids for these clusters using the mean value of each feature. This distribution of the data allows us to draw some conclusions about what features can best distinguish between types of users as categorized by this K-Means analysis. We also compared these clusters to the distribution of biology and non-STEM majors and the distribution of simple and complex analyzers. This allowed us to compare to both a relevant existing demographic marker as well as our key outcome metric.

3 RESULTS

3.1 Full Feature Set

Our full feature set was used to generate our first cluster model. This approach generated four clusters as defined in Table 2. In evaluating the centroids of each cluster, we see a low activity cluster (A), an iteration cluster (B), an information-seeking cluster (C), and a testing cluster (D). We observed the low activity cluster (A) is made up primarily of non-STEM majors and the high testing cluster (D) has a significantly higher ratio of biology majors. This pattern was aligned with the number of simple and complex investigations. We also noted that there were roughly equivalent numbers of majors and non-STEM majors in clusters B and C.

Table 2: Clusters Derived from Full Feature Set

	Feature	Cluster			
		A	B	C	D
Centroid	Actions Before H	0.1	2.3	7.6	0.3
	Frequency of I	0.1	0.4	0.6	0.0
	Frequency of H	0.3	0.2	0.1	0.2
	Frequency of T	0.4	0.3	0.2	0.7
	Frequency of C	0.2	0.1	0.1	0.1
	Number of I	0.3	3.9	8.9	0.3
	Number of H	1.2	1.5	1.2	1.3
	Number of T	2.1	3.0	2.3	5.9
	Number of C	1.0	1.1	1.0	1.0
	Number of I→T	0.0	0.5	0.3	0.0
	Number of T→I	0.0	0.0	0.0	0.0
	Maximum of Repeated T	1.9	2.2	2.0	5.3
	Maximum of Repeated I	0.3	2.9	8.4	0.3
	Transitions After Initial H	2.4	4.1	2.9	2.6
# of	Biology Major	10	7	7	18
	Non-STEM Major	29	9	15	10
	Simple Investigation	25	7	8	2
	Complex Investigation	11	6	12	24

3.2 Iteration Feature Set

To determine which of the features from our full feature set were most salient at determining clustering, we chose several primary features based on previous qualitative work [6], namely that experts tended to iteratively move back and forth between testing and pursue more information, which has

been observed both in SCI [6] and in novice/expert studies in engineering [2]. We use these features to explore differences between our subpopulations of novices. The iteration feature set generated four clusters which can be found in Table 3. In our iteration feature set, we observed very similar patterns as in our full data set, with cluster A (low activity) students more likely to be non-STEM major students with simple investigations. Meanwhile, cluster D students are more likely to be biology majors with high testing behavior and generally complex investigations. Clusters B and C remain a middle ground with relatively equivalent numbers of non-STEM and biology majors, along with roughly equivalent numbers of simple and complex investigations.

Table 3: Clusters Derived from Iteration Feature Set

	Features	Cluster			
		A	B	C	D
Centroid	Maximum of Repeated T	1.9	2.0	1.9	5.3
	Maximum of Repeated I	0.5	3.1	9.9	0.3
	Transitions After Initial H	2.3	4.7	2.4	2.4
# of	Biology Major	10	8	5	19
	Non-STEM Major	30	12	11	10
	Simple Investigation	26	8	6	2
	Complex Investigation	12	8	9	24

3.2 Frequency Analysis

Since our observed connections between clusters and simple and complex investigation designation could be related to frequency of actions performed, we next looked at the frequency of each action type (Table 4). In this table, the features selected for clustering analysis where the frequency of each type of action (H, I, T, and C) which prior work shows could be an indication of the level of expertise among subjects. Here, we see an interesting cluster of complex investigations among the users who did relatively more testing compared to other actions, whereas the number of information-seeking actions was not as critical. Additionally, we see that the individuals who focused the most on generating intermediate hypotheses were more likely to pursue simple investigations.

Table 4: Clusters Derived from Frequencies Feature Set

	Feature	Cluster			
		A	B	C	D
Centroid	Frequency of I	0.2	0.6	0.0	0.0
	Frequency of H	0.2	0.1	0.2	0.4
	Frequency of T	0.4	0.2	0.6	0.3
	Frequency of C	0.2	0.1	0.2	0.3
	Biology Major	2	12	24	4
# of	Non-STEM Major	4	27	22	10
	Simple Investigation	2	17	13	10
	Complex Investigation	2	17	31	3

4 DISCUSSION

Student understanding of practices is a key component of science education reforms in the United States [13]. Practices

such as inquiry may reflect student beliefs about the nature and process of science and how science knowledge is generated, or epistemological beliefs about science [17]. There is some suggestion that learning analytics techniques can be valuable for capturing and automatically assessing student practices [16] and potentially may be a better, more sensitive, and high throughput alternative to standard pen and paper assessments that are currently used to assess epistemological beliefs and science and/or nature of science understanding. We extend this work by demonstrating that we can cluster groups of novices based on their inquiry practices in a simulated authentic science inquiry experience. Our results provide quantitative evidence that corroborates with previous qualitative work in SCI simulations [17]. We found that abstract clickstream data aligns not only with previous qualitative work, but with a deeper analysis of overall investigative approach. We will now discuss our results in further detail with connection to possible future pedagogical approaches.

4.1 Alignment of clickstream data with previous qualitative work

Previous work [12, 17] suggested that experts tend to perform more information-seeking actions than novices, iterate back and forth between looking for information and performing tests, and perform complex investigations. We noted that transitions between information-seeking and testing were key features for determining clusters. We also noted that decision to perform long strings of tests or information-seeking was also a predominant characteristic of our clusters. We found that participants fell into four distinct clusters that can be described based on their predominant features. In Cluster A of our full feature analysis, we noted that these students performed roughly an average number of separate tests, were on the lower end for information-seeking actions, and did an average number of transitions between the two. We see this cluster consists primarily of non-STEM and the participants whose investigative strategies were simple in nature. This is in contrast to cluster D where we find more testing, but an equivalent amount of information-seeking actions and transitions to cluster A. We also noted in cluster A that these students tended to be perform many types of actions (e.g. generated more hypotheses), but were unable to pull each piece of their investigation together in a meaningful manner which may be why they were also likely to be coded as having simple investigations. This could reflect a lack of confidence in the material, which we discuss in further detail below. Cluster D is also predominantly biology majors and investigations that were coded as complex in nature. We also noted that complex investigations were not necessarily due to the frequency of actions performed (Table 4). For instance, performing lots of test (such as in Cluster C) did not necessarily lead to more individuals with complex investigations. We also found that major was weakly

correlated with performing a complex investigation ($\Phi = 0.33$, Cramer's $V = 0.24$, $p = 0.024$). Previous work suggests that non-STEM majors are more likely to hold misconceptions about how science works than biology majors [7], therefore our observation that biology majors are slightly more likely to perform complex investigations than non-STEM majors, where complex investigations are more expert like [17], supports the hypothesis that understanding of the process of science is reflected in practices. However, since the investigation type and relative experiences are weakly correlated, there are likely additional factors that are important for fully understanding the diversity of practices among these two populations of novices. Although predominantly information-seeking actions and the presence of iterations were also expert-like in nature, we did not observe this in Cluster D. Instead, these approaches predominated in cluster B (iteration) and C (information-seeking) which were not strikingly biology major or non-STEM major in nature. One possible explanation for this observation is that each of the clusters represent different facets of expert-like behavior, some of which (such as the use of testing) may be a stronger indication of understanding both of science practices and how practices are reflective of epistemological beliefs. Clusters alone may also be insufficient to properly understand the full spectrum of what it means to be expert-like in practices, and how these practices reflect underlying epistemological beliefs about science. Therefore, future work will connect clusters to other practices known to be connected to underlying epistemological beliefs about science, such as language [16]. Affect is another important construct to examine. Previous work [18] suggested that students who perform complex investigations also have higher intrinsic goal orientation, have higher self-efficacy in their ability to do science, increased motivation to do science, and are also more likely to identify as a science person. Therefore, further work is needed to holistically understand what these clusters mean in terms of more (or less) expert-like behavior, and how these clusters reflect understanding of science practices versus other factors.

4.2 Potential Pedagogical Interventions

In both our non-STEM major and biology major population we see students from each group falling into each cluster. This is consistent with previous qualitative work [17] where novices could be more-or-less similar to experts. Differentiation of groups of novices may be a possible avenue for classroom intervention and effectively connecting learning analytics to pedagogy, which is an ultimate goal of the learning analytics cycle [5] and an identified gap in the learning analytics community [9, 11, 14]. If an iterating student from cluster B was paired with an information seeker from cluster C in an authentic scenario, such as during a CURE or SCI simulation, how would their practices change? What about pairing a student from cluster D with one from cluster

A? Another strategy may be to use the information in the clusters to adapt pedagogical instruction. For example, say a class of students tends to fall into cluster A. Perhaps the instructor could encourage students to use information as part of their investigations, and also use this as an opportunity to discuss how investigations and science knowledge adapt and change in light of new evidence, which is both an important part of authentic science practice and a central tenant of both nature of science understanding and epistemological beliefs about science.

4.3 Limitations and Future Directions

One possible limitation of this work is that the expert behaviors we based our analysis on were conducted using a different version of the simulation, with different content and interface. Furthermore, our expert and novice populations in the original study resided in a different part of the USA. It may be possible that expert practices in the previous version of the simulation may be different in the newer version of the simulation, or that content or region differences may be impacting the data. Future work will analyze expert practices in the *Invasion of the Grackles* simulation using experts in the same geographic region as the novices reported here. These clusters are also based on clickstream data alone which we would predict would be only one of many important practices that are reflective of epistemological beliefs about science. Our future work will connect these clusters with novice performance on established metrics of nature of epistemological beliefs about science metrics in attempt to better understand what these clusters mean in terms of reflecting underlying epistemological beliefs about science and how these clusters relate to metrics like language [16].

5 CONCLUSIONS

In conclusion, our results suggest that clickstream analysis of inquiry practices is sufficient for distinguishing different groups of novices, namely non-STEM and biology majors. Future work will connect these findings with other relevant science practices that are also thought to be reflective of underlying epistemological beliefs about science, such as language [16], as well as affective factors such as self-efficacy and motivation. This work may indicate possible future pedagogical applications and approaches and underscores the prospective utility of using learning analytics in the classroom to assess difficult to measure constructs such as epistemology.

ACKNOWLEDGMENTS

We thank Emmy Royse, Jessie Sutton, Emily Wong and Shanna Smith for assistance with data collection and analysis.

REFERENCES

- [1] Abd-El-Khalick, F. 2012. *Examining the sources for our understanding about science: Enduring confluences and critical issues in research on nature of science in science education*. International Journal of Science Education, 34(3), 353-374.
- [2] Atman, C., Adams, R. Cardella, M., Turns, J., Mosborg, S., and Saleem, J. 2007. *Engineering design processes: A comparison of students and expert practitioners*. Journal of Engineering Education, 96(4), 359-379.
- [3] Auchincloss, L., Lauresen, S., Branchaw, J., Eagan, K., Graham, M., Hanauer, D., & Dolan, E. 2014. *Assessment of course-based undergraduate research experiences: A meeting report*. CBE Life Sciences Education, 13(1), 29-40.
- [4] Chinn, C. and Malhotra, B. 2002. *Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks*. Science Education, 86(2), 175-218.
- [5] Clow, D. 2012. *The learning analytics cycle: closing the loop effectively*. In Proceedings on the 2ns international conference on learning analytics and knowledge. 134-138. ACM.
- [6] Corwin, L., Graham, M., & Dolan, E. 2015. *Modeling course-based undergraduate research experiences: An agenda for future research and evaluation*. CBE Life Sciences Education, 14(1), es1.
- [7] Cotner, S., Thompson, S., and Wright, R. 2017. *Do Biology Majors Really Differ from Non-STEM majors?* CBE Life Sciences Education, 16(3), ar48.
- [8] Deng, F., Chen, D., Tsai, C., & Chai, C. 2011. *Students' views of the nature of science: A critical review of research*. Science Education, 95(6), 961-999.
- [9] Ferguson, R. and Clow, D. 2017. *Where is the evidence? A call to action for learning analytics*. In Proceedings of the Seventh International Learning Analytics and Knowledge Conference. 56-65. ACM.
- [10] Frank, E., Hall, M., and Witten I. 2016. *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"*, Morgan Kaufmann.
- [11] Hackbarth, A. 2017. *Are We Losing Sight of the Trees for the Forest? A Case for Localized Longitudinal Analytics*. In Shehata S, & Tan, J (Eds) Practitioner Track Proceedings of the 7th International Learning Analytics and Knowledge conference (LAK 17) 77-82. ACM.
- [12] Ketchen, D. and Shook, C. 1996. *The Application of Cluster Analysis in Strategic Management Research: An Analysis and Critique*. Strategic Management Journal, 17(6), 441-458.
- [13] National Research Council. 2012. *A Framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press.
- [14] Peffer, M. 2018. *Combining Multimodal Learning Analytics with Backward Design to Assess Learning*. Workshop Proceedings of the International Learning Analytics and Knowledge Conference.
- [15] Peffer, M.E., Beckler, M.L. Schunn, C., Renken, M., and Revak, A. 2015. *Science classroom inquiry (SCI) simulations: a novel method to scaffold science learning*. PLoS One, 10(3), e0120638.
- [16] Peffer, M. and Kyle, K. 2017. *Assessment of Language in Authentic Science Inquiry Reveals Putative Differences in Epistemology*. Proceedings of the Seventh International Learning Analytics & Knowledge Conference, 138-142.
- [17] Peffer, M., and Ramezani, N. (In Press) *Assessing Epistemological Beliefs of Experts and Novices via Practices in Authentic Science Inquiry*. International Journal of STEM Education.
- [18] Peffer, M., Royse, E., and Abelein, H. 2018. *Influence of Affective Factors on Practices in Simulated Authentic Science Inquiry*. International Conference of the Learning Sciences.
- [19] Quigley, D., Ostwald, J., and Sumner, T. 2017. *Scientific modeling: using learning analytics to examine student practices and classroom variation*. In Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17). ACM, New York, NY, USA, 329-338.
- [20] Sandoval, W. 2014. *Science education's need for a theory of epistemological development*. Science Education, 98, 3. 383-387.
- [21] Sandoval, W. 2005. *Understanding students' practical epistemologies and their influence on learning through inquiry*. Science Education, 89(4), 634-656.
- [22] Sandoval, W. and Redman E. 2015. *The contextual nature of scientists' views of theories, experimentation, and their coordination*. Science & Education, 24(9), 1079-1102.
- [23] Schizas, D., Psillos, D., and Stamou, G. 2016. *Nature of science or nature of the sciences?* Science Education, 100(4) 706-733.
- [24] Schwartz, R. and Lederman, N. 2008. *What scientists say: Scientists' views of nature of science and relation to science context*. International Journal of Science Education, 30(6), 727-771.