Cover Page

Educational Research in the 21st Century: Leveraging Big Data within Educational Research to Explore Teachers' Professional Behavior and Educational Resources Accessed within Pinterest

Dr. Kaitlin T. Torphy, Michigan State University, 620 Farm Ln., Rm. 240, East Lansing, MI 48824 torphyka@msu.edu

Hamid Karimi, Michigan State University, 428 Shaw Ln. Rm. 2333, East Lansing, MI 48824 karimiha@msu.edu

Dr. Jiliang Tang, Michigan State University, 428 Shaw Ln. Rm. 3115, East Lansing, MI 48824 tangjili@msu.edu

Dr. Kenneth A. Frank, Michigan State University, 620 Farm Ln., Rm. 462, East Lansing, MI 48824 kenfrank@msu.edu

Dr. Sihua Hu, Northwestern University, 2120 Campus Dr., Rm. 230, Evanston, IL 60208 sihua.hu@northwestern.edu

Research reported in this paper was supported by the Center for Business and Social Analytics at MSU, the National Science Foundation, and the William T. Grant Foundation under award numbers (NSF REAL– 1420532, WT Grant - 182764)

Thank you to the instructional content team including, Brianna Canedo, Emily Sinclair Emrick, Kim Jansen, Amanda Opperman, and Branda Peck.

Big Data and Educational Resources Within Pinterest

Abstract

Pinterest, a virtual resource pool and facilitator of social networks and social capital within virtual spaces is one example of a new resource that has the potential to provide instructional resources and professional support to teachers and inform educational policy. However, in an era of big data metrics, researchers must find meaningful approaches to characterize resources accessed and shared. This study investigates how we can leverage machine learning to characterize educational resources at scale within Pinterest—an image based platform.

Big Data and Educational Resources Within Pinterest

1. Introduction

The 21st century has seen individuals' unparalleled ability to retrieve information quickly, with little transactional cost. Acting as a virtual rolodex, the internet provides individuals opportunities to find and access resources, goods, and one another. Yet, with significant increases in informational volume has come greater necessity to efficiently seek out and procure that worth knowing (Fuchs, 2017; Bhaskar, 2016). This paper conducts an application of educational research using big data within social media to examine the quality of resources curated within teachers' mathematics instructional planning.

Given the potential of resources available online, people are increasingly engaging within social media online for their personal, social, and professional needs. By 2016, nearly 70% of a total 3.5 billion internet users reported using social media sites, such as Facebook and Pinterest.[1] Social media platforms represent unique niches within the virtual space that allow people to form communities and diffuse information across geographical and cultural boundaries. Virtual ties—those relationships established between individuals within social media—allow one to exchange resources and advice across a network of like-minded individuals. Networks within social media may provide individuals a heuristic when seeking out particular forms of advice or information (Fuchs, 2017). Rather than beginning in a vast pool of internet search, individuals may pose questions or request help in a smaller network of virtual ties within social media. These growing trends are expected to persist as internet technology evolves.

In the field of education, social media platforms have greatly expanded and transformed teachers' traditional social networks within the schoolhouse (Wellman, 2001). This novel social network in virtual space influences teachers' behaviors and exchange of information and

[1] https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

knowledge within their professional community of practices online (Hu & Torphy, 2018). Teachers may engage with one another and other invested individuals surrounding education policy change (Supovitz, Kolouch, Daly, & Del Fresno, 2017). They may turn to one another for information and instructional resources (Torphy, Hu, Liu, & Chen, 2017). The virtual ties they establish may begin within their schools, result from a recommendation from a social media algorithm, or reflect a teachers' particular professional orientation as they engage in online professional learning networks (Krutka & Carpenter, 2015). Furthermore, the knowledge and practices acquired from virtual resource pools (VRPs), those spaces in which teachers may access a large array of instructional resources, may be observed in classroom teaching and planning (Hu, Torphy, & Liu, 2018). In fact, preliminary research on teachers' planning as reflected within social media found the quality of instructional resources significantly related to teachers' quality of mathematics instruction in classrooms (Hu, Torphy, & Liu, 2018). Therefore, as education policymakers, researchers, and school leaders attempt to better understand and support the mechanisms that drive student achievement, it is worth considering how resources accessed online contribute to classroom instruction and academic success. That is, how teachers behave within social media may be a window into their conceptualization of their practice and instruction. Their professional reflections, accessed resources, and networks may allow for a more complete vision of 21st century teaching, at scale.

2. Prior research educational quality

In an era of accountability for students, teachers, and schools, social scientists and educational researchers have endeavored to identify what makes good schools, good (Lawrence-Lightfoot, 2008). Thus far, research indicates student and parental background contribute to the majority of differences in academic achievement (Hanushek, 2016). However, across all other school inputs, teacher quality remains the leading lever in positive change for student success (Wright, Horn, & Sanders, 1997; Rowan, Chiang, & Miller, 1997; Rivkin, Hanushek, & Kain, 2005).

Approaches to estimating teacher quality

Research on teacher quality spans across qualitative and quantitative approaches. Quantitative analysis often relies on state administrative datasets or other secondary datasets that follow a nationally representative sample of students or teachers over time (Figlio, Karbownik, & Salvanes, 2015). Qualitative approaches may examine the experiences of a set of teachers (Kennedy, 2005) through interview or survey to assess their practices, beliefs, and skills (Frank, Zhao, Penuel, Ellefson, & Porter, 2011). Other research has found students with teachers that had higher academic performance in college have better test scores. Yet, teacher quality extends beyond tacit knowledge to beliefs. Those teachers who focus on growth mindset, i.e. that learning is malleable, show higher student achievement than their peers (Dweck, 2007). Though teachers' academic aptitude and orientation to students' propensity to learn relates to their educational effectiveness, overall research indicates that teachers struggle with similar problems (Dweck, 2007; Lortie, 1975). Thus, teacher quality matters, yet it remains unclear what about teachers matters most and how teacher quality is affected by professional learning.

If teaching is a complex endeavor, educational effectiveness is a multifarious labyrinth comprised of nation and state, school and community, student and teacher. As we examine how teachers use resources we consider the multiple layers in which teachers work is embedded (Barr & Dreeben, 1977; Honig, Copland, Rainey, Lorton, & Newton, 2010). How teachers think frames their instructional planning and teaching enactment. Furthermore, how students respond frames teachers' future thinking and planning (Clark & Petersen, 1986). Subject content, available resources, school and teacher accountability pressure, district reform, and community engagement may impact educational effectiveness (Frank et al., 2013; Creemers & Kyriakides, 2007; Coburn, 2004; Anderson, 1986; Lowyck, 1980).

Research has found how teachers consider instruction and their students' needs relate to their effectiveness in the classroom (Palardy & Rumberger, 2008; Clark & Petersen, 1986). Furthermore, a large literature has explored the impacts of various curriculum on student gains (e.g. Boston & Wolf, 2006; Riordan & Noyce, 2001). In fact, the Common Core State Standards, a national curriculum reform, was premised on the notion that the tasks with which students engage, matter (Boston & Wolf, 2006). However, as teachers increasingly engage within social media and virtual space, they more autonomously direct the trajectory of their curriculum and classroom (Torphy, Hu, Liu, & Chen, 2017).

Teachers' Curation of Instructional Resources as a Data Artifact

Generally, examining impacts of teacher dispositions or planning on teaching has been costly—often approached through survey or observational studies of teachers' classroom enactment—and difficult to generalize (Salganik, 2018; Kane & Staiger, 2012). These large efforts require a non-negligible amount of resources to enlist participation, achieve high participant response rates, and physically interact through interview or classroom observation, with teacher participants. Furthermore, participating teachers may be impacted by observer bias, as they modify practices during evaluation of their instruction or in responses to surveys or interviews.

Social media affords the opportunity to observe individuals' thoughts and behaviors in real time, over time. Teachers in social media engage within various platforms (e.g. Twitter, Facebook, Pinterest, YouTube) for various purposes, from discussing educational policy to accessing supplemental instructional resources. Examining teachers' engagement within social media may provide an open window into their thought processes regarding instruction. How teachers think reflects their enactment of instruction. "What teachers do is directed in no small measure by what they think…it will be necessary for any innovations in the context, practices, and technology of teaching to be mediated through the minds and motives of teachers" (National Institute in Teaching, 1975, p. 1). Examining resources teachers curate for later reference allows a clearer analysis of the mechanisms underlying educational effectiveness. Each resource accessed, online space one follows, tweet posted, or virtual tie established contributes to an image of a teacher within 21st century schooling.

How teachers curate instructional resources may reflect their conceptualization of subject matter or teaching. Clark & Petersen (1986) reported a majority of teachers' planning is done mentally. More recently, a National Science Foundation Study of Elementary Mathematics Instruction examined teachers' mathematics instruction, lesson planning, and school-based collegial networks across four Midwestern states, and found a majority of teachers informally plan their mathematics lessons, often taking mental, rather than detailed written, notes (Frank et al., 2017). Today, the resources teachers choose to access and share within social media provide an observable artifact representing how teachers view their profession and work (Timperley & Robinson, 2001).

Curation is the beginning of a larger process of lesson development. "Elaboration, investigation, and adaptation are the phases through which teachers formulate their plans" (Clark & Petersen, 1986, p. 43). Online, teachers may visualize other teachers' enactment of lessons and instruction, outside their own schoolhouse bounds. Research shows visualization of instruction is an imperative component to instructional planning (Clark & Petersen, 1986). In a 21st century classroom, teachers' visualization may take many forms, be it through blogs, online social media, or personal experience. Thus, as more educators engage with one another online, their curation and shared understanding may relate to instructional relevance and sustainability (Liu, Torphy, Hu, & Tang, 2017).

3. Data and Methods

Utilizing big data analysis, we may view a window into teachers' professional lives as they plan and conceptualize their teaching practices. Bridging social science and computer science we may employ machine learning algorithms to scale understandings of what is being shared within education, from whom, to whom, and its relation to the changing education policy context. This work may provide deeper understanding of teachers' collaboration, conceptualization of instruction, and their profession providing richer information surrounding the question of teacher quality.

Pinterest

Pinterest, an image-based, personalized social media platform, which draws a population of 150 million active users per month, is a common social media platform and VRP that U.S. teachers use frequently for professional purposes. According to a national survey conducted by RAND, a majority of elementary and secondary teachers turned to Pinterest and its linked secondary sources to create and share instructional resources in response to a recent national education reform (Common Core State Standards Reform) or their own instructional needs (Opfer, Kaufman, & Thompson, 2016). Within Pinterest, teachers may encounter a personalized newsfeed of resources, from educational tasks to home improvement. Figure 1 illustrates an example Pinterest newsfeed.

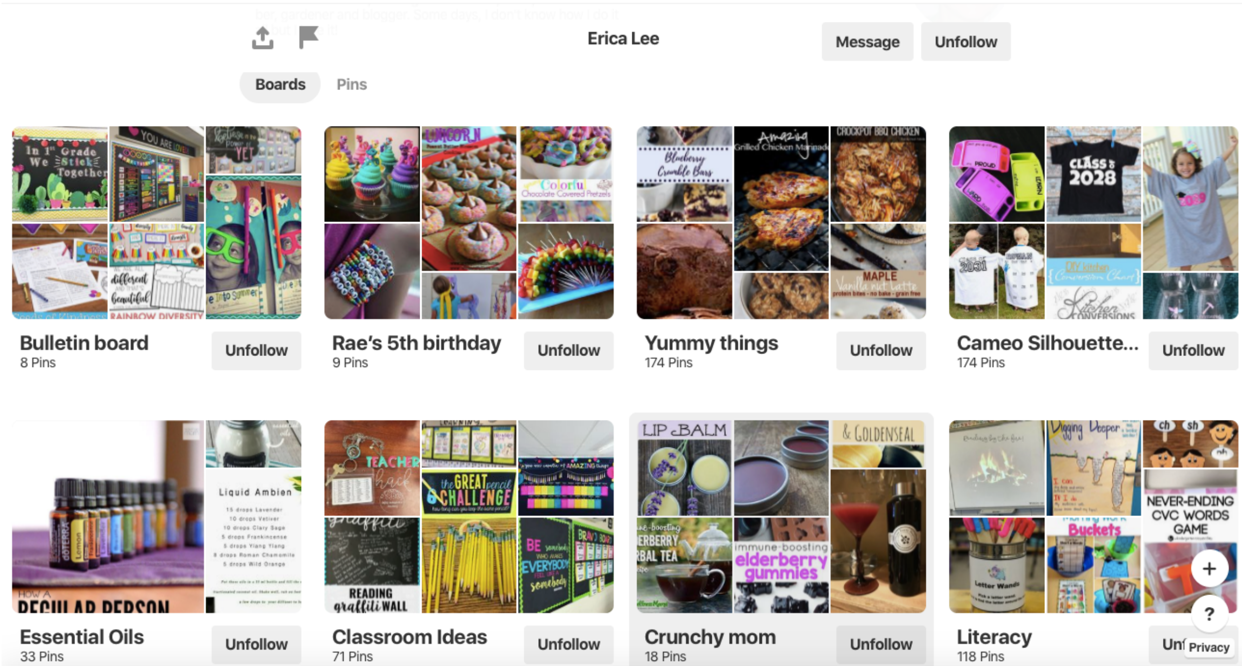
Big Data and Educational Resources Within Pinterest

Figure 1. An example Pinterest newsfeed

In Figure 1, images illustrate resources individuals may choose to access and save for later recall. Curating a particular resource may signal a teachers' desire to remember a specific resource for teaching and learning. Teachers curate resources across boards within Pinterest, depicted in Figure 2 below.

Big Data and Educational Resources Within Pinterest

Figure 2. An example teacher's Pinterest page



In Figure 2, a teacher curates professional and personal resources across self-created boards within Pinterest. A "message" button facilitates individuals ability to privately communicate with one another within Pinterest. An "unfollow" button allows individuals to follow or unfollow one another. As a follower or followee, teachers may connect with other professionals, creating virtual social networks—beyond the boundaries of the schoolhouse—engaging both professionally and socially. Teachers' informal and formal engagement within VRPs may facilitate the diffusion of teaching ideas and materials (Liu, Torphy, Hu, & Tang, 2017).

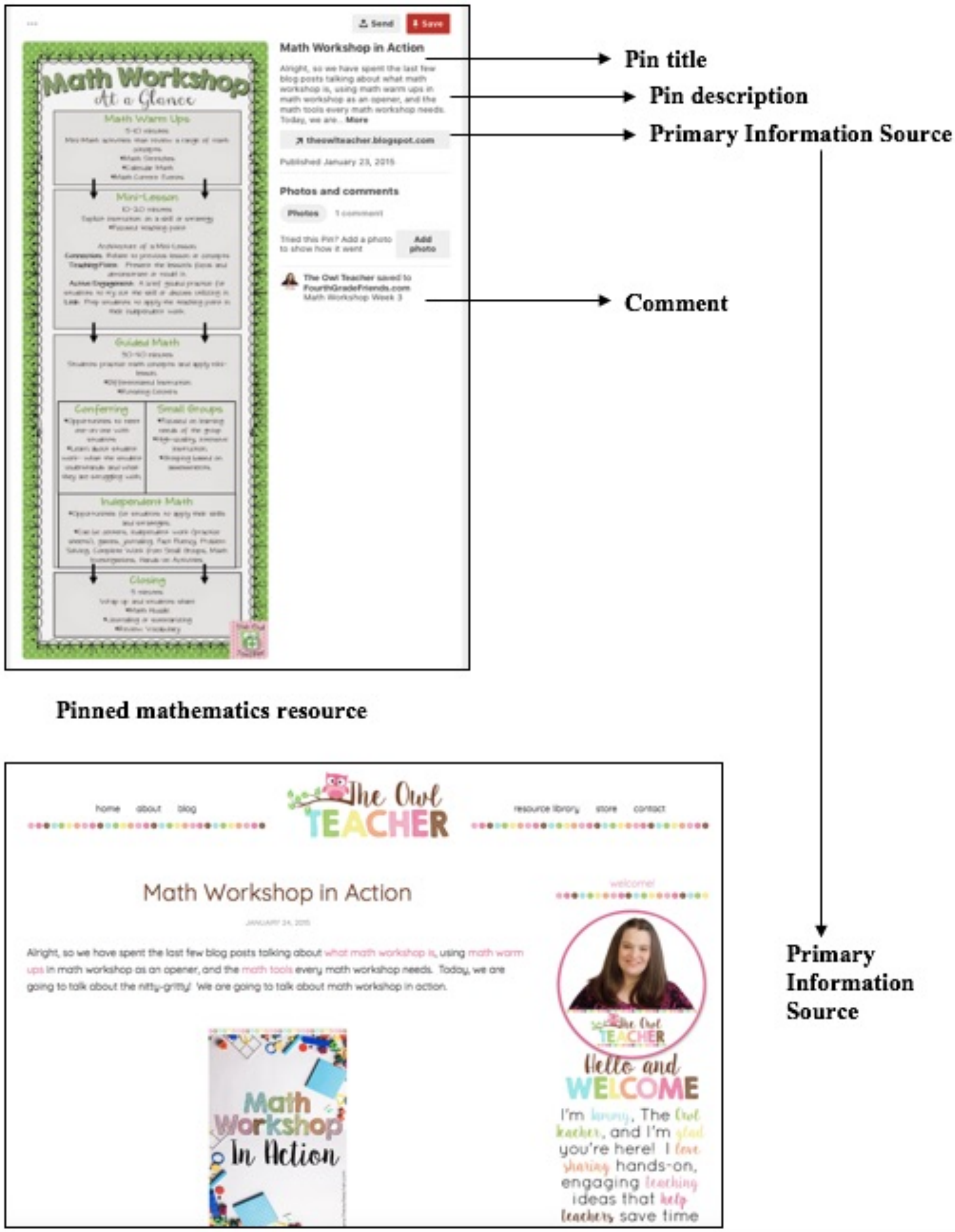
Despite teachers' prevalent use of social media for learning and teaching, there is little research on the mechanisms behind teachers' resource-seeking behaviors and the nature of the resources teachers access and share within and across virtual professional communities. According to a recent review of educational research related to social media sites, existing literature from 2004-2014 primarily emphasize college students' engagement with virtual spaces for learning outside of the classroom (Greenhow & Askari, 2017). Additionally, of work on education in social media, research has focused on virtual social networks and individual's online discourse, such as within Facebook (Manca & Ranieri, 2013; Manca & Ranieri, 2016) and Twitter (Gao, Luo, Zhang, 2012). An emerging body of studies has begun to investigate teachers' use of Pinterest to supplement curriculum and find varying degrees of quality for those resources residing within Pinterest (Hertel & Wessman-Enzinger, 2017; Hu, Torphy, Jansen, Opperman, & Lo, 2018). Still, there is a general lack of understanding regarding this educational phenomenon amongst K-12 teachers and the influence of knowledge acquired within VRPs and virtual communities.

The instructional resources teachers present to students mediate how students interact with academic content, one another, and their own learning. Yet, as teachers increasingly direct the resources utilized within their classroom, the vetting of curriculum may become diluted. Therefore, we seek to examine and characterize the instructional resources accessed and shared within Pinterest and, build an automated classifier to extend quality vetting at scale.

Characterizing the Quality of Mathematics Resources within Pinterest

We begin by examining the mathematics resources teachers curate within Pinterest. We use social media data to characterize the professional behaviors of teachers online, creating a training set to scale classification of teachers and mathematics content accessed and shared. We conceptualize teachers' engagement as indicants of their professional orientation, knowledge, and participation within an online professional learning community. Though this engagement may differ across virtual and physical spaces, we consider all teachers using social media for professional purposes committed and deliberate practitioners. Therefore, as teachers plan their instruction we expect they may generally follow pin links back to their primary source as they consider re-pinning a particular resource. Thus, pins present a window into the nature of teacher learning and knowledge acquisition. Figure 3 represents the connection between an example instructional resource “pinned,” that is accessed and saved, and the original source of information within the virtual universe. The resource pinned includes a title describing the pin, a longer, text description of the pinned resource, and a link to the origin of information—often outside Pinterest—within the virtual universe. If individuals repin a particular resource, they may comment on the resource and post pictures of their own adaptation within their classroom.

Figure 3. Instructional resource diffusion within Pinterest

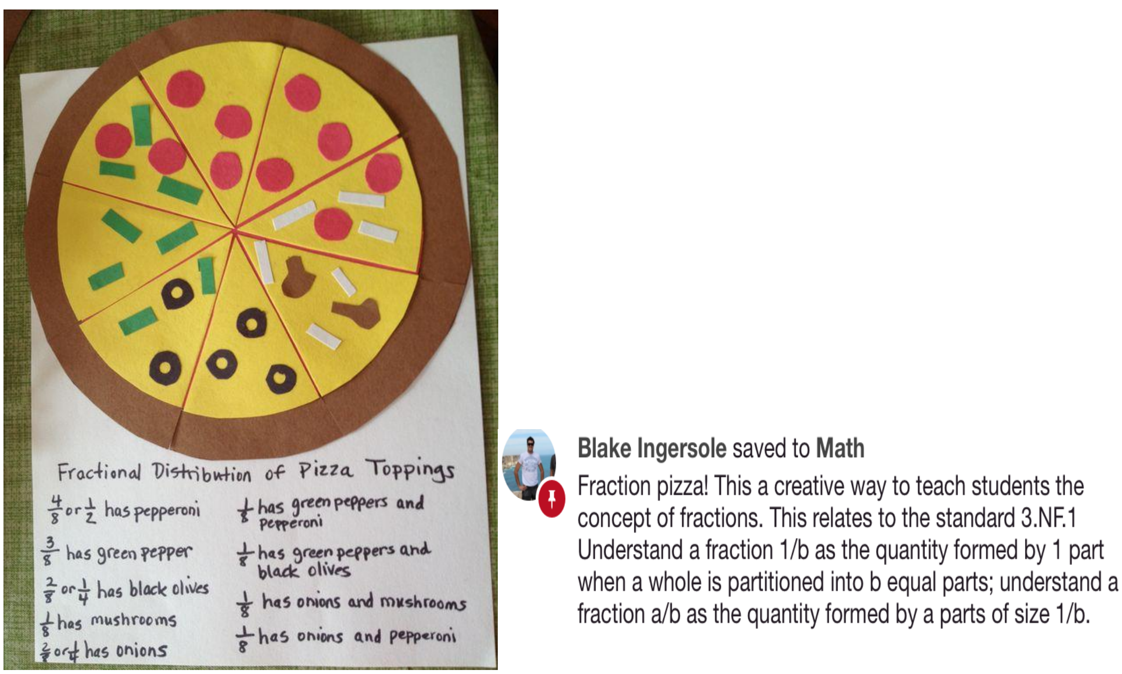


Big Data and Educational Resources Within Pinterest

Creating a dataset for classification

To examine and characterize mathematics resources within Pinterest, we assess the mathematical content, description, and linked source by pin, using the Revised Bloom's Taxonomy framework for human thinking order (Anderson & Krathwohl et al, 2001). The framework features six hierarchical cognitive levels required in human problem solving: remembering, understanding, applying, analyzing, evaluating, and creating. Thus, as we characterize the instructional content within Pinterest we seek to identify what function does the instructional content pinned serve and what does the mathematical task require from students. Figure 4 below illustrates an example of a mathematics task within Pinterest and a teacher's description of the pin.

Figure 4. Mathematics task and teacher's description of the intended student work within Pinterest.



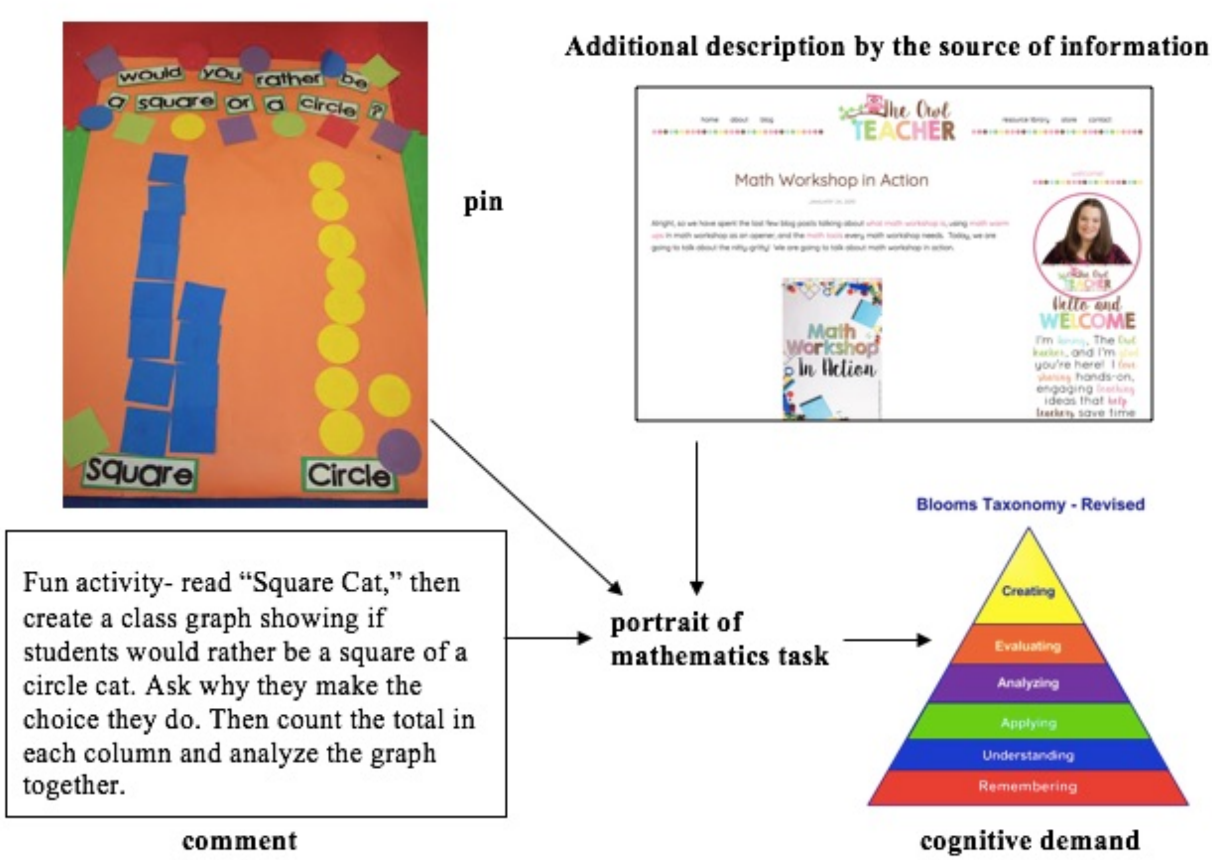
Both separately and collectively these indicants describe the quality and nature of instructional content accessed by teachers within social media communities. By characterizing the nature of instructional tasks accessed and shared we may better understand how teachers vary in their professional practices.

Data set

We conduct analyses of the instructional content with a sample of teachers from a NSF funded research project. Participants are in their first four years of teaching across four Midwestern states, namely Illinois (37 teachers), Indiana (65 teachers), Michigan (41 teachers), and Ohio (one teacher) over 4 years, from 2012-2016.

Hand Labeled Dataset. We compose an educational content team of four raters to code an initial 26 teachers' mathematical pins for cognitive demand (n=1,875 pins). For those pins with multiple tasks inherent in the pin, the team coded the primary nature of the task. All the pins were coded, at minimum, by two raters with an initial inter-rater agreement of 82%. Discrepancies were resolved to reach 100%. Figure 5 below presents the process to build a hand labeled dataset.

Figure 5. Building a hand labeled dataset for supervised learning



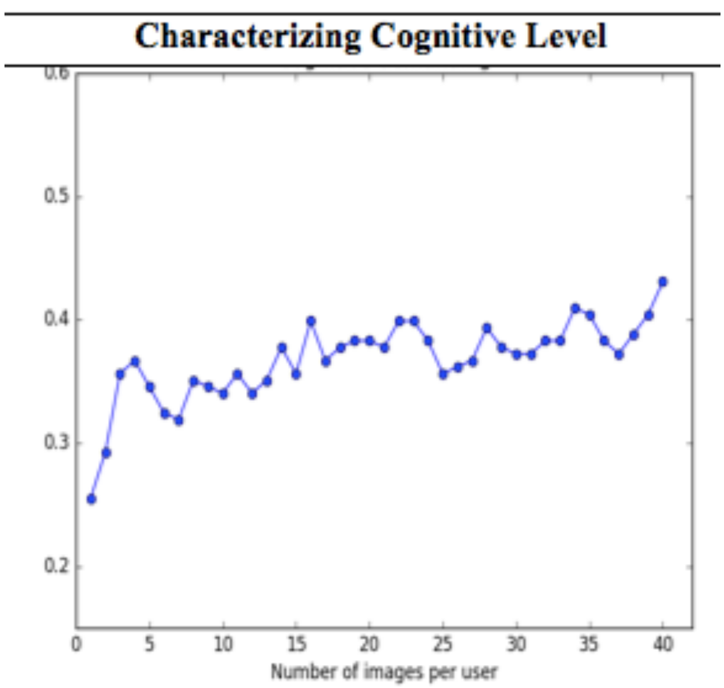
In Figure 5, a mathematics task depicts a student generated graph of circles and squares reflecting students' preference to be a square or circle cat. We use both the comment accompanying the pin, and—if there is not enough information to identify cognitive demand through image and pin—follow the pin's link back to the original source of information, outside Pinterest. In this task, students are required to organize and interpret simple data. Therefore, the task is labeled "analyze" and given a code of 4 for potential cognitive demand (see Appendix A for descriptions and examples of the potential cognitive demand of an instructional resource).

Big Data and Educational Resources Within Pinterest

We extended our sample to encompass an additional 119 teachers. Given that teachers vary in their pinning of educational resources, we chose to randomly sample mathematics pins to present an approximate portrait of a teachers' conceptualization of mathematics.

Identifying a threshold for random sampling. Using the initially coded data (n=1875 pins) we ran analyses using a convolutional neural network. Our results showed an unbalanced labeled data set, with a majority of instructional resources at the lower levels of cognitive demand. Given the unbalanced nature of our labeled dataset, we explored the number of pins needed, by teacher, to improve classification performance. We found increasing the number of images per teacher was helpful up to a particular threshold. See Figure 6 below.

Figure 6. Threshold for classification of mathematics resource cognitive demand within teacher



As shown in the left curve, the performance level for classifying the cognitive demand level is relatively stable after 15 pins, and 20-25 pins per teacher are sufficient to achieve desirable performance with our best available approach. That is, though teachers may continue to pin additional mathematics resources over time, their relative similarity in cognitive demand is stable. This result may be explained by literature on the stability of human behaviors over time.

Random sampling and labeling of mathematics resources. We extend our hand labeled dataset to encompass 119 teachers' mathematics resources curated within Pinterest. This results in 4853 resources pinned and hand labeled across years 2012-2016. We construct a stratified random sampling framework within teacher, across year. We find years vary in their pinning volume, with 2012 being the greatest in resources curated. Therefore, we oversample 2012 and 2016 to account for the additional volume within 2012 and balance data in the last year of observations. Table 1 below presents the sampling framework by year for a given teacher.

Table 1. Stratified random sampling framework for hand-labeled coding of cognitive demand by teacher, over year

Year	Number of Mathematics Tasks Pinned
2012	15
2013	5
2014	5
2015	5
2016	15

Table 1 presents the maximum number of mathematics resources classified and hand labeled for a given teacher in our extended training set. Some teachers may not have pinned 45 mathematics resources across all years. In these cases, we labeled all teachers' mathematics tasks pinned. In contrast, other teachers may have joined Pinterest after 2012, or not pinned mathematics resources in a particular year. For these teachers, we sampled a commensurate number of instructional pins across all other years, while still oversampling in 2016, to equal the maximum 45 mathematical resources pinned.

Feature Extraction

After we code the pins, we utilize them to train a pin classifier. Figure 7 illustrates the feature extraction and classification processes. Since cognitive demand levels 4, 5, and 6 are very close semantically and we do not have a sufficient number of pins within each category, we aggregated them into a single level of cognitive demand i.e., class 4. Therefore, as shown in Figure 7, we perform a 5-class classification where class 1, class 2, and class 3 correspond to cognitive demand levels 1-remembering, 2-understanding, and 3-applying, respectively. Class 4 represents cognitive demand levels 4-analyze, 5-evaluate, and 6-create altogether. Class 0 represents those pins that do not fall into any of cognitive demand levels i.e., they are not applicable. These pins may be an organizational resource, a pedagogical approach to teaching, or a conglomerated packet of instructional resources. In any case, they are not a single instructional task and therefore cannot be coded for a mutually exclusive classification of cognitive demand.

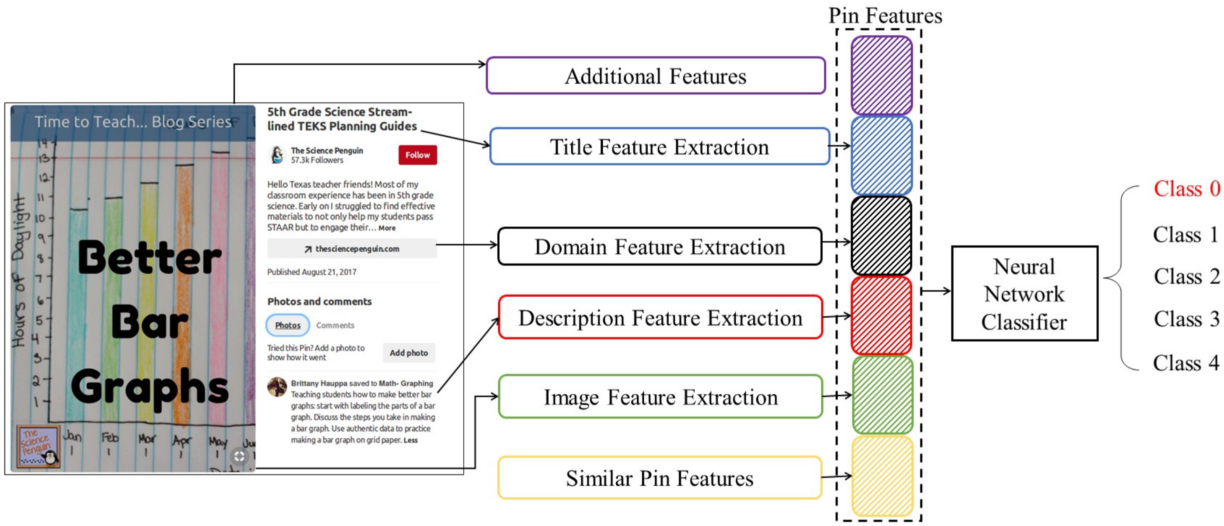


Figure 7. Pin classification and feature extraction

To perform the pin classification, we extracted a set of features from pins. We extracted features from six sources of data described as follows:

Image Feature Extraction. We extracted visual features related to images of pins via a pre-trained neural network (inception-v3 from google[2]). This network has been trained on a a large data known as *Imagenet*[3] that includes more than 14 million images from various categories of images. The pre-trained network using inception-v3 yielded a continuous-values vector of size 512 for an input image. Thus, utilizing the *pre-trained* network, we extracted a vector of size 512 for each pin's image encoding visual features.

Description Feature Extraction. Descriptions are another source of data which provide useful information about pinned resources. We trained a statistical topic modeling know as Latent Dirichlet Allocation (LDA) introduced by Blei et al. (2003). We transfer descriptions of a

2 (https://arxiv.org/abs/1512.00567)
3 (http://www.image-net.org/)

pin to a vector of size 10. Each element of that vector indicates the probability (score) of the description belonging to a particular topic identified by LDA. Note that if the description of a pin is not available, we set the vector to zero. See Appendix C, Table C1 for topics and word groupings.

Title Feature Extraction. Pin title may also help to identify the cognitive demand level of that pinned resource. Similar to description, we utilized LDA and transferred each title to vector of size 10 representing scores of the title belonging to extracted topics. Note that if the title of a pin is not available, we set the vector to zero. To review topics and word groups, see Appendix C, Table C2.

Domain Feature Extraction. Pins in Pinterest link to various Internet sources e.g., an education blog, a nonprofit organization, etc. In an effort to increase the performance of pin classification, we took into account the domain (URL) from which the pins originate. Building on work from Torphy, Hu, Liu, and Chen (2017), we recognized eight different categories for domains and labeled more 16,000 domains to these categories. The domain feature used for our pin classification is a one-hot vector of size 9 where one element corresponding to the category a pin's domain is 1 and the rest is zero. If no category was available for a pin's domain, the 9th element of the one-hot vector is set 1 and the rest will be zero.

Additional features. Other than the basic features described above, we included three additional sets of features. The first set includes the U.S. state associated with the pinner (i.e., the teacher who has pinned the pin). We represented state features of a pin with a one-hot vector of size 4 corresponding to four states Illinois, Indiana, Michigan, and Ohio, respectively. We considered teachers with a greater number of followers may pin resources of a particular quality (i.e. category). Therefore, we included a second set of additional features, the number of followers and followees by teacher, by pin. We called this set of features *Following*. Finally, not all resources pinned within Pinterest are free, some link back to original sources that monetize their resources for sale. We considered the potential that resources may differ in their quality dependent on their cost and therefore included a "For Profit" label.

Similar Pin Features. On the one hand, coding cognitive demand is a time-consuming and demanding process. One the other hand, for a better classification performance more data is needed. To add additional analytical power, increase classification accuracy, and extend coding in an efficient approach, we connected hand labeled pins to those most visually similar within Pinterest. For each pin within Pinterest, the algorithm recommends a number of visually similar pins. We connected to the first two pins recommended amongst a set of similar pins suggested by Pinterest. We connected to only two pins after examining the threshold at which classification accuracy across hand labeled and visually similar pins drops off. The threshold analysis experiment is included in Appendix B. For each coded pin, we extracted features described above (except additional features, given the fact that these pins extend outside our sampled pinners) from its corresponding two similar pins and added them to the features of the coded pin.

Classification

Once we extracted features for all 4,853 pins and all of their similar pins, the data was ready to perform classification. We randomly selected 80% of pins from each class for training and the remaining 20% for testing. Table 2 presents the details of our data set statistics. We used a neural network for a supervised classification. This network takes features of a pin and tries to refine the features in a way to map them to the provided cognitive demand levels inherent within a pin e.g., cognitive demand 3 in Figure 7. After fully training the neural network, we used it to classify pins in the test set (unseen samples).

Table 2. Descriptive statistics for hand labeled data set

Cognitive Demand Level (Sample class)	**Train Set**	**Added Similar Pins**	**Test Set**
0	1376	2752	343
1	763	1526	190
2	1050	2100	262
3	508	1016	126
4	188	376	47
Total	**3885**	**7770**	**968**
	11655		**12623**

4. Results

Evaluation metrics. We measured the performance of pin classification on the test set via F1-score. F1-score is the harmonic mean of precision and recall defined as F1-score = 1/(1/precision + 1/recall) and precision= TP/(TP+FP) and recall = TP/(TP+FN) where FT, FP, FN are the number true positive, false positive, and false negative samples (pins). The range of F1-score is between 0 and 1 and the close this metric to 1, the better the performance[4].

[4] Refer to https://en.wikipedia.org/wiki/Precision_and_recall for more explanation and discussion about F1-score in data mining

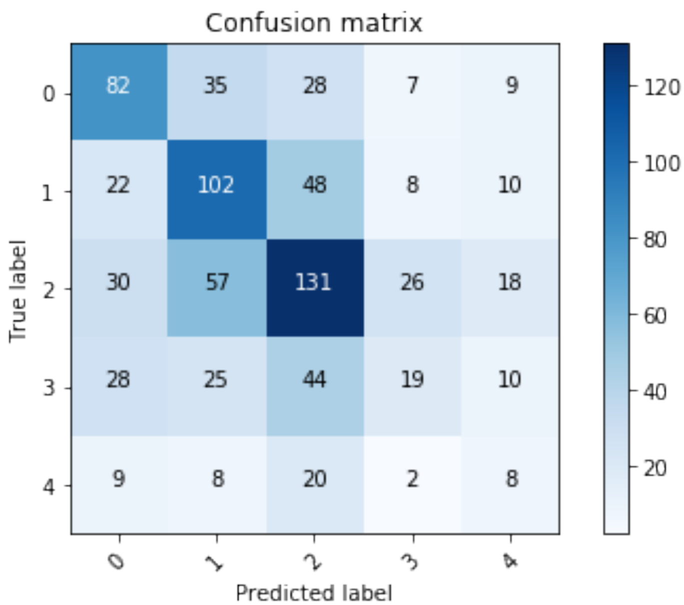
Table 3. Pin Classification Results. Best performance has been shown in bold

NO	Included Features (Sources)	F1-Score
1	Image, Similar pins	0.38
2	Image, Description, Similar pins	0.38
3	Image, Description, Domain, Similar pins	0.38
4	Image, Description, Domain, State, Similar pins	0.39
5	Image, Description, Domain, State, Title, Similar pins	0.35
6	Image, Description, Domain, State, Title, Following, Similar pins	0.41
7	Image, Description, Domain, State, Title, Following, For Profit Site, Similar pins	0.41
8	**Image, Description, Domain, Title, Following, For Profit, Similar pins**	**0.43**
9	Image, Description, Domain, Title, Following, For Profit	0.38
10	Random guess	0.24

We performed pin classification on all possible feature subsets i.e., conducting 255 (2^8-1) experiments since we have 8 different feature sets (see Figure 7). Table 3 illustrates the results of several experiments. Based on the results in Table 3, we found including multiple feature sets generally improved the performance. This suggests our labeled feature sets complement one another in pin classification for cognitive demand. Also by comparing the performance of experiment No.8 and No.9, we observed improvement in classification when features of similar pins were added. This suggests adding similar pins for those most conservative cases, i.e. when the pin is nearly identical, can significantly improve the classification performance. Finally, though we achieved only 43% accuracy, in all experiments we significantly outperformed random guess. This indicated that our pin classification algorithm indeed learns discriminatory and useful information about the semantics of cognitive demand levels reflected in feature sets.

To better understand the results, we included the confusion matrix of the experiment No.8 (i.e., the top performing experiment) in Figure 8.

Figure 8. The confusion matrix



The confusion matrix is by and large diagonal. This indicated that our pin classifier correctly classified pin's cognitive demand the majority of the time. As shown in Figure 8, errors are predominately made across adjacent categories (e.g., sometimes 2's are coded as 3's). This is encouraging because adjacent categories are semantically close, and therefore, mislabeled classification deviate less from the true value. For example, the pinned resource depicting a graph of squares and circles in Figure 5, labeled as "*analyzing* (value 4) requires students to organize and interpret data. It could have been classified as *applying* (value 3), carrying out or using a procedure to conduct a one-step problem. However, the confusion matrix suggests it is unlikely that pinned resource will be misclassified as *remembering* (value 1)--retrieving knowledge or recognizing facts. For those misclassified values, we noted the confusion matrix looks symmetric (sometimes 2 are coded as 3's, sometimes as 1's). This suggests misclassification is likely related to random measurement error.

Practical Applications of a Classifier for Cognitive Demand within Pinterest

Given the social nature of Pinterest, the resources individuals pin are widely and virally shared across the entire network. This is particularly true for education pins where teachers embrace sharing and re-pinning instructional resources within their own communities. This brings about a useful and quite practical application for our pin classifier i.e., identifying the cognitive demand levels of those pins that are identical to our hand labeled sample. In fact, we evaluated this capability and achieved a high Fl-score of 0.67. This is over three times the accuracy of a random guess for cognitive demand classification, and nearly 1.5 times the accuracy of classification results within the most conservative estimation trials in which we ensured the test set had no identical resources. Thus, for practical application within virtual space, we may consider leveraging our developed pin classifier to characterize both those unknown resources and visually identical resources therefore significantly increasing the proportion of pins able to be closely classified at-scale.

5. Discussion

This work contributes to big data analysis within education, leveraging computer science tools to characterize and classify the instructional resources teachers curate within Pinterest, a prevalent social media platform within education. Through these efforts, we provide early work on the state-of-the-field as teachers navigate their curriculum and seek out supplemental resources for their classroom. We find that teachers pin relatively consistent quality resources within Pinterest, as the marginal return to neural network classification past approximately 25 pins diminishes. Given the field's emphasis on lifelong learning and ongoing professional learning (Darling-Hammond et al., 2009), it is interesting that the types and quality of resources teachers access is relatively fixed. This may be due to the distribution of instructional resources available online if those resources are predominately low quality. However, it may also echo education literature that finds teachers, in general, replicate their own learning experiences (Lortie, 1975; Kennedy, 2005) and are not, in general, risk takers (Howard, 2013). Therefore, for example accessing tasks that require students to memorize or apply mathematical concepts may be more conservative and similar to their own lived experiences as a student.

Scaling Instructional Quality Classification within Pinterest

Attempts to scale classification of mathematics instructional tasks show moderate success in conservative testing experiments but greater capacity to characterize instruction quality in the Pinterest platform. As many instructional resources are curated, being pinned and re-pinned within the social network, classifying those most popular (prevalent) pins may allow for research based quality ratings across a wide range of resources. Leveraging visual similarity indices and those identical or nearly identical pins, we may better be able to identify quality instructional resources at scale.

In conservative testing experiments, we find increasing the features considered in classification improves accuracy, overall. Combining features of the pin image, pin description, pin title, linked outside domain, for profit orientation of linked domain, number of individuals a teacher follows, and visually similar pins accurately classifies cognitive demand 43% of the time. These results suggest researchers should use a variety of approaches to characterize educational resources within Pinterest. As complex social media data, we expect various attributes from image features, to text descriptions, to teachers' social networks may improve classification. Future work will incorporate teachers' social network as predictive of pins' instructional quality.

Implications for Education Research

This work combines qualitative approaches to computer science applications in machine learning to characterize mathematics instruction and scale analysis of educational resources within virtual resource pools. Examining the types of resources accessed and shared online may provide deeper insight into teachers' instructional planning and decision-making. Furthermore, classifying educational resources at scale may provide the capacity to quickly identify high and low quality resources. This information could be paramount to teachers and administrators as they seek supplemental resources that meet the needs of their students and local context. This is one of many potential applications of big data within social media applied to problems of educational practice.

Social media data may be viewed as educational artifacts curated by teachers over time (Torphy, Hu, Liu, & Chen, 2017). A window into teachers' professional lives, it is updated in real time, iteratively. Moreover, as a prevalent social media resource in which a majority of teachers report engagement (Opfer, Kaufman, Thompson, 2016), ignoring diffusion of resources, information, and teachers' professional learning within virtual space may obscure understandings of reform enactment, teaching, and learning within 21st century schools.

Challenges

Characterizing semantic content and potential cognitive demand within instructional resources is a difficult task. Across a team of education expert raters, maintaining inter-rater reliability and alignment to the Revised Bloom's Taxonomy requires regular check-ins and discussion. Utilizing convolutional neural networks to characterize these images requires finding patterns in complex images and identifying the inherent cognitive demand, often within a photo of student work or a classroom. Across human and computer assisted approaches, identifying quality within the complexity of education images proves challenging.

Our sampling suggests the distribution of instructional resource quality within Pinterest is unbalanced, with a majority of mathematics resources reflecting lower cognitive demand. The unbalanced data provides less opportunity to train and test our classifier across higher categories of cognitive demand, therefore we find less accuracy as we encounter those higher quality resources.

Future Work. In this work, we simply used the number of followers and followees of teachers as network features i.e., we included “node degrees” in network analysis. However, we plan to extend our analysis of network diffusion, utilizing the network beyond just simple follower/followee counts. In particular, the communication between teachers reflected in re-pinning, commenting, etc. may help increase the accuracy of resource classification. Furthermore, via advanced network features (e.g., local clustering coefficient) and network analysis approaches (e.g., network embedding), we can develop practical instructional applications and extend current algorithm practices, recommending a teacher to another one if they the share behaviors in curating same resource quality, or recommending virtual mentors who excel at curating instructional resources within Pinterest.

6. Conclusion

Social media in education is omnipresent and continues to expand its reach (Education Week, 2018). How teachers connect to one another, and the sources of information they choose to access and share may contribute to understandings of social capital inequity and variation in teacher quality. In kind, big data may complement administrative and deeper qualitative data gained from interviews and survey. Educational research may gain opportunities to examine, at scale, how teachers respond to education policy, their students' needs, and their collegial social networks—both within and outside the schoolhouse.

Big data derived from social media may be applied to a variety of educational problems of practice ranging from impacts of reform to teaching behaviors in the classroom. Engaging a multi-disciplinary team of social scientists and computer scientists may broaden how we attempt to measure and facilitate professional learning within this sphere. This first attempt to develop a machine learning algorithm to characterize mathematics instructional resources is one example of a greater opportunity to examine teachers' behavior in an ongoing way and at a scale that holds potential to improve the policies, resources, and opportunities available to educators.

References

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., ... & Wittrock, M. C. (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives, abridged edition. *White Plains, NY: Longman*.

Anderson, L. W. (1986). Research on Teaching and Educational Effectiveness. *Curriculum report*, *15*(4), n4.

Barr, R., & Dreeben, R. (1977). 3: Instruction in Classrooms. *Review of research in education*, *5*(1), 89-162.

Bhaskar, M. (2016). Curation: The power of selection in a world of excess. Hachette UK.

Boston, M., & Wolf, M.K. (2006). Assessing academic rigor in mathematics instruction: The development of the Instructional quality assessment toolkit. CSE Technical Report 672. Center for the Study of Evaluation, Standards and Student Testing, UCLA.

Clark, C., & Peterson, P. (1986). Teachers' thought processes. In M. C. Wittrock (Ed.), *Handbook of Research on Teaching* (3rd ed., pp. 255-296). New York: Macmillan.

Creemers, B., & Kyriakides, L. (2007). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge.

Coburn, C. E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of education*, *77*(3), 211-244.

Darling-Hammond, L. Wei, R., Andree, A., Richardson, N., & Orphanos, S. (2009). Professional learning in the learning profession. *Washington, D.C.: National Staff Development Council.*

Dotterer, A. M., & Lowe, K. (2011). Classroom context, school engagement, and academic achievement in early adolescence. *Journal of Youth and Adolescence, 40*(12), 1649-1660.

Dweck, C.S. (2007). *Mindset: The New Psychology of Success.* New York: Penguin Random House LLC.

Education Week. (2018, January 10). Data: Where do teachers get their ideas? *10 Big Ideas.* Retrieved from https://www.edweek.org/ew/section/multimedia/data-where-do-teachers-get-their-ideas.html?cmp=eml-en-eu-news2&M=58363164&U=1932747

Figlio, D., Karbownik, K., & Salvanes, K. G. (2015). Education research and administrative data. *Handbook of the Economics of Education, 5.*

Frank, K. A., Bieda, K., Salloum, S., Ye, Y., Hu, S., & Youngs, P. (2017). *Brass tacks and networks: The network effects of substantive planning on beginning elementary teachers' ambitious mathematics instruction (Paper in symposium).* The 2017 Annual Meeting of the American Educational Research Association (AERA 2017), San Antonio, T.X.

Frank, K.A., Zhao, Y., Penuel, W., Ellefson, N., & Porter, S. (2011). Focus, fiddle and friends: Experiences that transform knowledge for the implementation of innovations. *Sociology of Education, 84*(2), 137-156.

Fuchs, C. (2017). Social media: A Critical Introduction. Sage.

Gao, F., Luo, T., Zhang, K. (2012). Tweeting for learning: A critical analysis of research on microblogging in education published 2008-2011. *British Journal of Educational Technology, 43*(5), 783-801.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology, 78*(6), 1360-1380.

Greenhow, C. & Askari, E. (2017). Learning and teaching with social network sites: A decade of research in K-12 related education. *Education and Information Technologies,* 1-23.

Hanushek, E. A. (2016). What matters for achievement: Updating Coleman on the influence of families and schools. *Education Next, 16*(2), 22-30.

Hertel, J. & Wessman-Enzinger, N. (2017). Examining Pinterest as a curriculum resources for negative integers: An initial investigation. *Education Sciences, 7*(45), 1-11.

Honig, M. I., Copland, M. A., Rainey, L., Lorton, J. A., & Newton, M. (2010). Central office transformation for district-wide teaching and learning improvement. *Center for the Study of Teaching and Policy, University of Washington*.

Howard, S.K. (2013). Risk-aversion: Understanding teachers' resistance to technology integration. *Technology, Pedagogy, and Education, 22*(3), 357-372.

Hu, S., Torphy, K., Jansen, K., Opperman, A., & Lo, Y. (2018). What do teachers share within socialized knowledge communities: A case of Pinterest. *Journal of Professional Capital and Community, 3(*2), 97-122.

Hu, S. & Torphy, K. (2018). Social media in education: Curation within and outside the schoolhouse. *Handbook on Social Media Analytics: Advances and Applications. In press.*

Kane, T. J., & Staiger, D. O. (2012). Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains. Research Paper. MET Project. *Bill & Melinda Gates Foundation.*

Kennedy, M. M. (2005). *Inside Teaching.* Cambridge, M.A.: Harvard University Press.

Krutka, D. G. & Carpenter, J.P. (2016). Participatory learning through social media: How and why social studies educators use Twitter. *Contemporary Issues in Technology and Teacher Education, 16*(1), 38-59.

Lawrence-Lightfoot, S. (2008). *The good high school: Portraits of character and culture*. Basic Books.

Lortie, D.C. (1975). *Schoolteacher.* Chicago: The University of Chicago Press.

Lowyck, J. (1980). Teaching skills and teacher education. *Revue ATEE Journal*, *3*(4), 173-190.

Manca, S. & Ranieri, M. (2013). Is it a tool suitable for learning? A critical review of the literature on Facebook as a technology enhanced learning environment. *Journal of Computer Assisted Learning, 29*(6), 487-590.

Manca, S. & Ranieri, M. (2016). Is Facebook still a suitable technology-enhanced learning environment? An updated critical review of the literature from 2012-2015. *Journal of Computer Assisted Learning, 32*(6), 503-528.

Opfer, V. D., Kaufman, J. H., & Thompson, L. E. (2016). Implementation of K-12 state standards for mathematics and English Language Arts and literacy. Santa Monica, CA:RAND Corporation.

Palardy, G. J., & Rumberger, R. W. (2008). Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis, 30*(2), 111-140.

Penuel, W. R., Frank, K. A., Sun, M., Kim, C. M., & Singleton, C. (2013). The organization as a filter of institutional diffusion. *Teachers College Record, 115*(1), 1-33.

Ragin, C. (1994). *Constructing social research: The unity and diversity of method.* Thousand Oaks: Pine Forge Press.

Riordan, J. E., & Noyce, P. E. (2001). The impact of two standards-based mathematics curricula on student achievement in Massachusetts. *Journal for Research in Mathematics Education*, 368-398.

Rivkin, S.G., Hanushek, E.A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.

Rowan, B. Chiang, F., & Miller, R. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education, 704*(4), 256-284.

Sanders, M. G. (1998). The effects of school, family, and community support on the academic achievement of African American adolescents. *Urban Education*, *33*(3), 385-409.

Supovitz, J. A., Kolouch, C., Daley, A.J., & Del Fresno, M. (2017). #Common Core project: How social media is changing the politics of education. Retrieved from http://repository.upenn.edu/hashtagcommoncore/3

Tang, J. & Liu, H. (2014). An unsupervised feature selection framework for social media data. *IEE Transactions on Knowledge and Data Engineering, 26*(12), 2914-2927.

Timperley, H. S., & Robinson, V. M. (2001). Achieving school improvement through challenging and changing teachers' schema. *Journal of Educational Change*, *2*(4), 281-300.

Torphy, K. Liu, Y., Hu, S., & Chen, Z. (2018). *Sources of professional support: Patterns of teachers' curation of instructional resources in social media. (Paper in symposium).* The 2018 Annual Meeting of the American Educational Research Association (AERA 2018), New York, N.Y.

Torphy, K., Hu, S., & Liu, Y. *The influence of social media on lesson preparation and practices (Paper in symposium).* The 2018 Annual Meeting of the American Educational Research Association (AERA 2018), New York, N.Y.

Big Data and Educational Resources Within Pinterest

Torphy, K., Hu, S., Liu, Y., & Chen, Z. (2017). Teachers turning to teachers: Teacherpreneurial behaviors within social media. *(Paper in symposium).* The 2018 Annual Meeting of the American Educational Research Association (AERA 2017), San Antonio, T.X.

Wellman, B. (2001). Computer networks as social networks. *Science, 293*(5537), 2031-2034.

Wright, S.P., Horn, S. P., & Sanders, W.L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11,* 57-76.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993-1022.

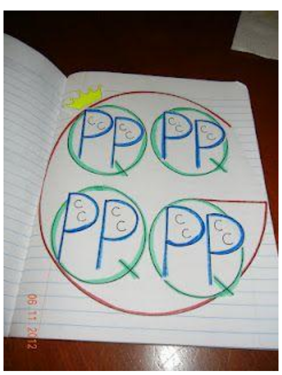
Appendix A

Definition and Images of Mathematics Cognitive Demand

We find the majority of resources are coded in the lower two levels of cognitive demand, which expect students to exercise thinking order to remember (481 pins) and understand (459 pins). Moreover, 205 resources are characterized at the apply level, with much less at the higher three levels. The unbalanced data distribution adds another layer of complexity to train neural network models as they signal the potential of a left-skewed true population of educational resource quality accessed by early career teachers within Pinterest. Below are example pins for each cognitive demand level and corresponding definitions.

Table A1. Definition and Images of Primary Mathematics Cognitive Demand Evidenced on Pinterest

Level 1: Remembering

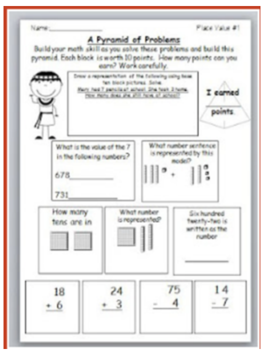


Recall mathematical facts and basic mathematical concepts. *Example verbs: Define, list, repeat, state, memorize, duplicate, recognize, recall*

Level 2: Understanding

Construct mathematical meaning from oral, written, and representational forms. *Example verbs: explain, classify, describe, identify, locate, recognize, report, infer, summarize, interpret*

Level 3: Applying



Use the mathematics knowledge in a different situation/context, or in a new way. *Example verbs: execute, demonstrate, illustrate, implement, solve, calculate*

Appendix B

Threshold analysis to determine number of similar pins

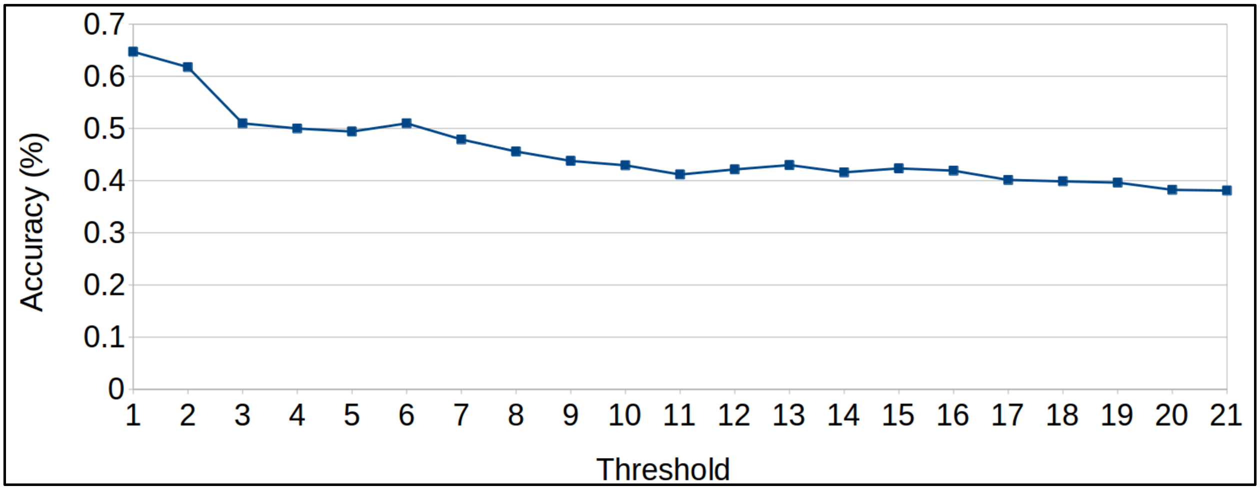


Figure B.1. Threshold analysis experiment

We experimented identifying the threshold to determine how many visually similar pins we can/should include to each hand labeled pinned resource. Therefore, we first collected 21 similar pin from Pinterest for 18 coded pins. We referred to 18 original hand labeled pins as "anchor pins." Then, we coded all 21 visually similar pins for each one of the anchor pins i.e., we identified the cognitive demand of 378 (18*21) visually similar pins. Note that similar pins for each anchor pin within Pinterest are ordered based on an internal algorithm for visual similarity. At each time, we added a single visually similar pin for every anchor pin and calculated the accuracy i.e., the number of added similar pins that have the same value with their corresponding anchor pins/total number of similar added pins. Figure B.1 illustrates the results of this experiment. As shown in this Figure, we found the more similar pins that are added, the more performance (i.e., accuracy) declined. We observed that by adding two visually similar pins per each anchor pin, we achieved an acceptable accuracy (> 60%) that those included similar pins had the same cognitive demand level as their anchor pins. Therefore, in the conducted experiments, we added two similar pins for each coded pin.

Appendix C

Latent Dirichlet Allocation of Pins Descriptions and Titles

Table C1. Topic Words by Pin Description

Topic	**Words**
Math Centers and Practice for Mastery	0.067*"game" + 0.051*"card" + 0.033*"math" + 0.027*"practic" + 0.023*"activ" + 0.023*"addit" + 0.022*"time" + 0.018*"task" + 0.018*"student" + 0.017*"center"
Holiday Themed Resources	0.039*"activ" + 0.025*"page" + 0.024*"color" + 0.022*"christma" + 0.021*"famili" + 0.018*"money" + 0.015*"word" + 0.014*"worksheet" + 0.013*"winter" + 0.012*"student"
Word Problems and Strategies for Solving	0.061*"student" + 0.028*"problem" + 0.019*"help" + 0.017*"multipl" + 0.017*"strategi" + 0.014*"read" + 0.014*"stori" + 0.013*"solv" + 0.012*"question" + 0.011*"word"
Free Resources and Activities	0.041*"place" + 0.035*"valu" + 0.029*"freebi" + 0.027*"free" + 0.022*"printabl" + 0.021*"count" + 0.021*"activ" + 0.016*"graph" + 0.016*"color" + 0.014*"kid"
Math Anchor Charts and Bulletin Boards	0.083*"chart" + 0.056*"anchor" + 0.035*"shape" + 0.029*"board" + 0.021*"school" + 0.013*"bulletin" + 0.012*"classroom" + 0.012*"math" + 0.012*"calendar" + 0.011*"craft"
Number Sense Practice and Resources	0.142*"number" + 0.044*"student" + 0.025*"practic" + 0.020*"page" + 0.018*"frame" + 0.016*"worksheet" + 0.016*"line" + 0.016*"languag" + 0.014*"includ" + 0.013*"sens"
Common Core and Daily Math Activities	0.131*"math" + 0.079*"grade" + 0.028*"center" + 0.019*"work" + 0.018*"review" + 0.018*"common" + 0.018*"core" + 0.015*"daili" + 0.015*"activ" + 0.014*"week"
School and Classroom	0.028*"classroom" + 0.028*"teacher" + 0.021*"book" +

	0.020*"student" + 0.019*"read" + 0.017*"school" + 0.015*"organ" + 0.013*"year" + 0.013*"idea" + 0.012*"great"
Common Core Materials	0.068*"grade" + 0.058*"common" + 0.057*"fraction" + 0.055*"core" + 0.032*"math" + 0.030*"standard" + 0.019*"measur" + 0.017*"googl" + 0.016*"kindergarten" + 0.016*"teach"
Interactive Notebook Lessons	0.119*"word" + 0.047*"write" + 0.038*"sight" + 0.024*"interact" + 0.024*"notebook" + 0.023*"lesson" + 0.018*"roll" + 0.016*"plan" + 0.015*"teach" + 0.015*"spell"

Note 1. The numbers are coefficients/score of words in a topic.

Note 2. Stem of words are shown.

Table C2. Topic Words for by Pin Title

Topic	**Words**
Early Elementary Math Resources	0.144*"freebi" + 0.132*"math" + 0.059*"kindergarten" + 0.043*"blog" + 0.041*"count" + 0.033*"core" + 0.032*"multipl" + 0.032*"game" + 0.031*"common" + 0.031*"teach"
Math Free Resources	0.166*"math" + 0.067*"grade" + 0.054*"free" + 0.051*"step" + 0.049*"lemon" + 0.048*"literaci" + 0.043*"fall" + 0.036*"kid" + 0.032*"kindergarten" + 0.031*"kinder"
Fractions, Addition, and Subtraction Activities	0.080*"school" + 0.070*"fraction" + 0.070*"addit" + 0.057*"learn" + 0.054*"activ" + 0.049*"freebi" + 0.048*"math" + 0.037*"educ" + 0.037*"subtract" + 0.032*"grade"
Place Value and Geometry Worksheets	0.177*"place" + 0.172*"valu" + 0.061*"teach" + 0.039*"shape" + 0.033*"printabl" + 0.027*"freebi" + 0.027*"worksheet" + 0.025*"learn" + 0.025*"number" + 0.024*"littl"
Math Games and Manipulatives	0.176*"math" + 0.088*"game" + 0.073*"grade" + 0.046*"build" + 0.039*"number" + 0.037*"resourc" + 0.036*"elementari" + 0.033*"teacher" + 0.030*"class" + 0.026*"teach"

Early Elementary Math	0.221*"math" + 0.113*"number" + 0.071*"grade" + 0.070*"kindergarten" + 0.037*"shape" + 0.034*"unit" + 0.033*"love" + 0.030*"learn" + 0.028*"sens" + 0.025*"workshop"
Math Centers and the Common Core	0.239*"grade" + 0.134*"math" + 0.061*"teach" + 0.042*"second" + 0.041*"center" + 0.034*"common" + 0.034*"core" + 0.032*"geometri" + 0.021*"money" + 0.018*"chart"
Math Games and Telling Time	0.175*"math" + 0.080*"time" + 0.078*"game" + 0.067*"teach" + 0.056*"learn" + 0.035*"tell" + 0.032*"free" + 0.026*"tidbit" + 0.026*"tunstal" + 0.025*"card"
Strategies for Charts and Graphs	0.078*"math" + 0.070*"graph" + 0.066*"chart" + 0.048*"strategi" + 0.045*"anchor" + 0.042*"addit" + 0.039*"subtract" + 0.038*"common" + 0.038*"core" + 0.034*"regroup"
Math Classroom Resources	0.090*"number" + 0.082*"math" + 0.078*"classroom" + 0.072*"free" + 0.056*"freebi" + 0.056*"teacher" + 0.046*"miss" + 0.044*"activ" + 0.043*"idea" + 0.039*"teach"

Note 1. The numbers are coefficients/score of words in a topic.

Note 2. Stem of words are shown.