The Influence of Teacher Created Metadata in Online Resource Exchanges

Samuel Abramovich
University of Pittsburgh
LRDC – Suite 830
Pittsburgh, PA, USA
1-412-324-3909
sja24@pitt.edu

Chris Schunn
University of Pittsburgh
LRDC – Room 821
Pittsburgh, PA, USA
1-412-624-8807
schunn@pitt.edu

ABSTRACT
Online resource exchanges offer a different paradigm for how teachers find and select resources. As more teachers choose to go online to gather resources, questions remain about what factors influence their selection of resources. Using decision heuristics theory as a lens, we created a hierarchical linear model from the dataset of an online teacher resource exchange for a national teaching organization. Our specific focus was to discover what teacher generated resource metadata predicts number of downloads. Based on our findings, there is little support to suggest that the majority of users rely on a simple heuristic. We also found that a high number of low ratings predict more downloads than a resource with a low number of high ratings. This seemingly runs counterintuitive to the idea that more low ratings would dissuade teachers from looking at a resource.

Categories and Subject Descriptors
K.3.1 [**Computer Uses in Education**]: Collaborative Learning

General Terms
Human Factors

Keywords
Education Resource Exchanges, Decision Heuristics, Hierarchical Linear Models

1. INTRODUCTION
Teachers are increasingly turning to online solutions to meet their professional needs [1]. For example, online teaching resource repository are considered a productivity booster for teacher lesson planning [2]. However, questions remain about what features of online resources can influence teacher selection. Prior research in lab studies of knowledge management systems suggest that the number of user-entered ratings does not affect content searches [3]. A survey conducted on online reviews in a non-education setting indicated that moderate reviews were more helpful than extremely positive or negative comments [4]. These types of conclusions could cause educational technology designers to alter their plans for creating and improving teacher resource exchanges. Yet little prior research has specifically targeted the unique aspects of large-scale online teacher resource exchanges. Because the evaluation of educational resources can have a direct effect on teacher practice and teacher learning [5-7], and consequently student performance [8], it is imperative that organizations that provide online teacher resource exchanges unpack how teachers find and select resources in these online systems.

1.1 Theoretical Background
The advantage of online teacher resource exchanges is that they both facilitate the sharing of resources through teacher contribution and also utilize user ratings and comments to act as assessment of the resources. Ostensibly, there would be little need for professional development or training in how to use the resources from an online teacher resource exchange since resources would have been created by teachers for other teachers. Resource ratings and comments can act as a quality filter and further enhance selection of the best resources since, theoretically, teachers should be accurate raters of quality instructional resources.

However, creating online teacher communities is also a complex process [9]. Johnson [10] argues that teacher communities of practice requires careful scaffolding in order to exist online. Morris & Hiebert [11] identify the potential for classroom teaching to improve through collaborative resource creation.

The unique needs of teachers must be accounted for in order to design teacher resource exchanges for maximum effectiveness. When finding an online learning resource, teachers often do not know or misapply the appropriate pedagogical practices that should accompany the resource [12]. The solution of teaching teachers how to effectively use online resources is also not ideal given the challenges of designing effective technology based professional development [13].

Many online teacher resource exchanges utilize simple resource metadata to enable optimal discovery of resources. This metadata, such as content type, grade level, difficulty level, and duration, is similar to that generated for online learning resources repositories [14]. However, there are several important differences between teaching resources and learning resources. Learning resources are any digital document that can be used for learning [15]. High-quality teaching resources might include some of the same information as a learning resource but should also include additional information. For example, math teachers have been known to take learning resources that are designed to induce high-level mathematical thinking and during instruction proceduralize the mathematics, lowering the effectiveness of the resource [16]. A properly designed teaching resource might not include a digital resource or artifact but instead could provide information on effective pedagogies, provide subject matter content for the teacher, or suggest ways that a teacher can relate learning to other parts of a curriculum [17].

To further help teachers filter through the abundance of resources available online on any given topic, teacher resource exchanges often utilize *evaluative metadata*, typically in the form of user ratings and comments. In order to develop a theory for how teachers use *evaluative metadata* in online resource exchanges, we began with a more general body of research on human decision-making processes.

Research on decision heuristics proposes the existence of a stopping rule [18]. In the case of an educator looking for a resource online, the stopping rule would be the heuristic that triggers the educator to stop looking at other resources, if just momentarily, and download the resource selected. A stopping rule might be simple or complex, perhaps depending on the teacher and their goals when searching for a resource. Todd and Gigerenzer [19] detail several simple heuristics for stopping rules that have accounted for human decision making in situations ranging from car purchases to college selection to voting (see Table 1). The *Recognition* heuristic suggests that, when using an online resource exchange, a resource would be selected if it resembles resources that were previously chosen. *Take the Best* is a one-rule heuristic that suggests that whatever singular decision factor appeared to be best in prior download decisions would be used singularly going forward. For example, if an individual thought a high rating was the best indicator for downloading a resource then they would use a high rating as their heuristic going forward. *Take the Last* is similar to *Take the Best* but instead of the best decision the last decision is given primary importance. If reading a comment on a resource was the last factor that swayed someone to download that resource then the same individual going forward would look at comments primarily as a heuristic for downloading a resource.

The described heuristics all have advantages and disadvantages. The *Recognition* heuristic would be efficient if an educator was an expert in recognizing high quality resources. However, novice teachers using this heuristic could be selecting resources they recognize from their limited experiences but are not optimal for their teaching needs. *Take the Best* and *Take the Last* could be very efficient in identifying resources for download but also cause an individual to download a resource they would have skipped using more complex or contextually sensitive heuristics.

More complicated heuristics for stopping rules include *Dawes' Rule,* which would suggest that the decision to download a resource would be based on comparing the number of positives (e.g. high ratings, positive comments) to the number of negatives (e.g. low ratings, negative comments). *Franklin's Rule* is a variant that uses weighted accounting of pros and cons. As an example, a resource's mean rating and number of comments may both be considered when choosing what to download, but number of comments could factor more heavily than the mean rating. The *Multiple Linear Regression* heuristic would suggest that the decision to download a resource is a weighting account of pros and cons but also taking into the strength/extent of the positives and negatives. For example, a resource with one positive rating would be evaluated for download differently that a similar resource that had an equal number of positive and negative ratings.

Table 1. Stopping Rule Heuristics

Heuristic	Explanation
Recognition	Based on a resemblance to prior chosen resources
Take the Best	Using the one best factor or test that had previously worked for selecting a resource
Take the Last	Using the one factor or test that had last worked for selecting a resource
Dawes' Rule	Counting the number of Pros vs. Cons
Franklin's Rule	A weighted counting of Pros vs. Cons
Multiple Linear Regression	A weighted sum of the extent of Pros vs. Cons

1.2 Data Source

We examined the data metrics from TFANet - Teach for America's (TFA) website that supports the exchange of teacher resources. TFA is a non-profit organization that seeks to take high-performing, recent college graduates and place them for two years in urban and rural schools in the US with underserved students. For the 2011-2012 school year, TFA will enlist 9000 corps members who will teach 600,000 students and is projected to continue to grow with the goal of placing 15,000 teachers in 60 different regions by 2015. In addition, TFA continues to receive both private and public funding, including a 2010 $50 million dollar grant from the US Federal Department of Education.

TFANet was created as a way to support corps members by providing an online network where a variety of resources and services can be accessed. Corps members in need of a teaching resource can go to the resource exchange on the TFANet website and search for a type of resource based on keywords and metadata about each resource. A TFANet user can perform a keyword search for a resource and pre-select by grade, subject, file type, resource, type, author, appropriateness for the school year, or state specificity. According to TFA, during the 2010 Fall semester, 75 percent of corps members downloaded resources from TFANet.

Each resource in TFANet also has a web page that supplies additional information about the resource. This information includes a detailed description provided by the author, an average rating of one to five stars from other TFANet users, the number of ratings, comments from other TFANet users, and a Blue Ribbon icon if TFANet administrators identified the resource as high-quality. For the rest of our analysis, we refer to this information as *evaluative metadata*.

The administration at TFA graciously provided us a copy of the main TFANet database from multiple time points. Included in the database are individual resource names, descriptions, dates of upload, ratings, number of comments, number of ratings, number of downloads, and author's name and region. Descriptions of TFANet's user interface are a result of direct interaction with the system. All of the information provided in this paper has been confirmed for accuracy with TFA administration.

2. METHODS

Our underlying hypothesis for studying online teacher resource exchanges is that *evaluative metadata* will influence a teacher's choice to download a resource. Thus a resource that is highly rated or rated many times would predict the eventual number of times the resource will be downloaded.

To investigate this general hypothesis, we examined the following more specific research questions:

RQ1. Do *evaluative metadata* in online teacher resource exchanges predict significant variance in the number of downloads?

RQ2. Which *evaluative metadata* are most predictive of downloads?

RQ3. Are download decisions best described by a simple decision heuristics like Take the Best, or more complex decision heuristics like Franklin's Rule or Linear regression?

On this last point, we are seeking to characterize the heuristics of groups of teachers to inform overall. Thus, we are not going to focus in this paper on identifying patterns that are individually true of all members nor can we suppose that simple individual decision strategies average to a generally accurate decision strategy [20].

For our analysis, we chose to examine data from TFANet between Feb. 10 and Mar. 10 of 2011. By February, most teachers have established a routine and are deep into instruction and February occurs before most traditional intense school testing periods in the US. For this one-month timeframe, there were 26,959 unique visitors to the resource exchange with 178,626 searches for resources and 79,348 downloads.

Our next step was to establish what *evaluative metadata* was available to users for deciding on whether to download a resource. We were able to identify 9 different variables visible to users, which are all listed in Table 2. Our list of metadata variables is not exhaustive but we believe we have selected the most influential of the metadata available, based on its availability to users and ease of understanding.

We chose a hierarchical linear model (HLM) analysis as an appropriate method to examine the data given both the design of the TFANet data and the hierarchical nature of resources in online teacher resource exchanges. Social data commonly has a nested nature, meaning that there is repeated observations of behaviors attributed to one person or group of people Ignoring the influence of these repeated observations violates the assumption of independence necessary for certain statistical analyses such as linear regression. HLMs do not require an assumption of independence for all data since its methodology considers cross-level effects (e.g. how different authors might vary in their creation of resources) and variance between levels of data (e.g. how much a resource's appeal is based on the author) [21].

Much like students in a classroom, resources from an author cannot be assumed to be independent of each other. Traditionally, teachers will write a lesson plan even if re-using a learning resource. The reason is that teacher professional development stresses the importance of altering instruction according to specific student needs. Yet, just as students in a classroom or school will have some level of homogeneity, so too will teaching resources from a particular author, whether it is from the author's content knowledge or teaching experience. Hierarchical linear models allows for data that has a nested structure. Consequently, HLM estimates of the standard errors will appropriately account for the nesting in our data [21].

We structured the TFANet data for analysis so that resources and their associated metadata were nested within individual authors. Our HLM models were created using HLM 7.0 from SSI Inc. [22] A resource's file format was converted to a binomial variable indicating whether the format was easily editable (e.g. Microsoft Word, Microsoft PowerPoint) or not easily editable (e.g. PDF, JPEG). Also converted to a binomial was whether the resource had a blue ribbon indication of quality and whether the author was a current corps member.

To include resource descriptions in our model without tackling the time intensive task of coding commentary, we created a variable of the number of characters used in each resource's description. Our assumption is that the length of a description can also represent the level of detail of the description (i.e. the longer the description the more detailed the description). While this is not a perfect solution, it does provide the advantage of avoiding unintended bias by overemphasizing content qualities of the resource descriptions. For example, a history teacher would benefit from a different type of resource description than a math teacher. Because our analysis covered all content areas, using the resource description length served as a fair measure for all subject areas.

Table 2. Descriptive Statistics

Variable	**Mean**	**Min**	**Max**
Resource Level (n=16,863)			
Dependent Variable			
Downloads during 1 Month	3.24	0	51
Independent Variables			
Character Count of Resource Descriptions	279	5	2477
Date of Upload (Age of Resource)	Jul. 11, 2009	Aug. 14, 2008	Feb. 9, 2011
Number of Ratings (Not Including Missing Ratings)	2.38	1	38
File Format is Editable (1=No)	0.11	0	1
Blue Ribbon Indicator (1=Yes)	0.06	0	1
Number of Comments	0.43	0	23
Missing Average Rating (1=Yes)	0.34	0	1
Average Rating between 1 and 2.99 (1=Yes)	0.08	0	1
Average Rating Between 3 and 3.99 (1=Yes)	0.20	0	1
Average Rating Between 4 and 4.99 (1=Yes)	0.28	0	1
Perfect 5.0 Rating (1=Yes)	0.10	0	1
Person Level Independent Variables (n=2,149)			
Author's Year in TFA	2007.2	1990	2010
Author Currently in TFA (1=No)	0.69	0	1

New resources can be supplied by current core members, alumni and various types of TFANet administrators. However, we removed administrator-generated resources because we wished to focus on the exchange of materials among teachers.

The rigor of the TFA admissions process results in a teacher population that is highly motivated. TFA also attracts individuals who are comfortable with self-starting and bootstrapping their own teaching education since TFA's mission is to place teachers in high-needs schools with less training than traditional teacher education programs. TFA members should also be fairly technically savvy since they are all recent college graduates and experienced with the current state of learning technologies used at higher-education institutions. Finally, TFA is a national organization, which means that its members represent a wide variety of organizational and policy contexts. By focusing on teacher-generated resources, our findings can provide insight to what many would consider the prototypical type of teacher who would benefit from participating in a national online resource exchange.

Our analysis was performed on 16,863 resources written by 2,149 different authors. Not all of the resources were rated, but that was to be expected based on the prior work examining TeachersPayTeachers.com [23]. Unrated resources are important to include in any analysis of a teacher resource exchange for two reasons. One is that they are naturally occurring and as a result must be considered when looking at overall system behaviors. Second is that a missing rating might increase the influence of other metadata. For instance, if a user is interested in a resource and sees that it is unrated then they might be more influenced by other metadata than for a resource where ratings are available. To include the missing data, we dummy coded average rating into five categories representing similar ratings with broadly equal distributions. Thus the coefficients from our findings are all compared to the norm of no rating.

The overall statistical model included all the variables shown in Table 2. Specifically, the model used in our analysis was:

$\eta_{ij} = \gamma_{00} + \gamma_{01}$*'*Author's Year in TFA*$_j$' + γ_{02}*'*Current TFA Member*$_j$' + γ_{10}*'*Number of Characters in Description*$_{ij}$' + γ_{20}*'*Date of Upload*$_{ij}$' + γ_{30}*'*Number of Ratings*$_{ij}$' + γ_{40}*'*File Format is Editable*$_{ij}$' + γ_{50}*'*Blue Ribbon Indicator*$_{ij}$' + γ_{60}*'*Average Rating between 1 and 2.99*$_{ij}$' + γ_{70}*'*Average Rating between 3 and 3.99*$_{ij}$' + γ_{80}*'*Average Rating between 4 and 4.99*$_{ij}$' + γ_{90}*'*Perfect 5.0 Rating*$_{ij}$' + γ_{100}*'*Number of Comments*$_{ij}$'+ u_{0j}

To calculate the predicted change in downloads by metadata, we chose as our dependent variable the number of downloads over a one-month period. In other words, we calculated the increase, if any, of the total number of downloads for each resource between Feb. 10th and Mar. 10th. Despite some of the resource metadata changing over course of the month, like number of ratings or average rating, we felt confident that a one month time frame beginning with when the metadata was recorded and ending with the increase in downloads was an appropriate starting point for our analysis. A much shorter time frame would have too few downloads to study, and a much longer time frame would greatly increase the occurrence of changing metadata over the studied download period.

The number of downloads is count data (i.e., only zero or positive integers with a particular skewed distribution) and, like other count data, is better fit by a Poisson distribution than a Normal distribution. Therefore our HLM model included a log-link function to account for this distribution of our outcome data. We used a constant exposure Poisson model in HLM 7.0 to understand whether any of the *evaluative metadata* could predict the number of downloads in a month for TFA member-created resources. We also added an estimate of over-dispersion to adjust for the dependence of the independent variables on the mean of the outcome variable.

3. FINDINGS

We ran a null model with no independent variable to determine whether nesting resources among authors could account for any variance in resource downloads. The results were significant and the calculated Odds Ratio of 3.21, the expected means of number of downloads for the month, is very close to the true mean of downloads 3.24. Consequently a multi-level model does seem to be a valid method of predicting resource downloads.

The estimates from the full model are shown in Table 3. Three variables were not found to be statistically significant and thus, given the very large N of this analysis, are unlikely to be used in a download heuristic: the date when the resource was uploaded, the number of comments about a resource, and the year the author of the resource was in TFA. As number of comments and number of ratings tends to be highly correlated (r=.749, p<.001), the number of comments may not add information.

Table 3. Full Model

Variable	**Notation**	**Coefficient**	**Odds Ratio**
Resource Level			
Base***	γ_{00}	1.137	3.117
Character Count***	γ_{10}	0.0002	1.000
Upload Date	γ_{20}	0.000	1.000
Number of Ratings***	γ_{30}	0.086	1.090
File Format is Editable*	γ_{40}	-0.088	0.915
Blue Ribbon Indicator***	γ_{50}	0.132	1.141
Average Rating between 1 and 2.99	γ_{60}	0.042	1.043
Average Rating Between 3 and 3.99***	γ_{70}	0.120	1.127
Average Rating Between 4 and 4.99***	γ_{80}	0.171	1.186
Perfect 5.0 Rating***	γ_{90}	0.293	1.341
Number of Comments	γ_{100}	-0.022	0.978
Person Level			
Corp Year in Teach for America	γ_{01}	-0.010	0.991
Current Corp Member or Alumni**	γ_{02}	-0.127	0.881

*p<.05, **p<.01, ***p<.001

The number of characters in the resource description was found to be significant (β=0.000164, p<.001). The small coefficient stems from relatively larger means; if shifting to how many hundreds of characters, the units would be similar to what was found for the effect of number of ratings on number of downloads. Thus, a simple decision heuristic based on this finding could be that educators decide to download a resource based on the approximate length of the resource description.

The editable qualities of the file-format, blue ribbon indication, and whether the author was identified as a current member of TFA were binomial variables and all found to be statistically significant. Any of the variables could also be used in a simple heuristic for choosing to download a resource. However, their binomial nature combined with the relatively small coefficients means they do not explain a lot of the variance in downloads. If these variables were used in a common simple heuristic then it would seem they should predict more of the variance. It might be that they have less weight in more complex decision rules or individuals rarely use them in simple decision rules.

The coefficients for the various ratings levels should be interpreted relative to the no rating case. In general, most rating levels produced more downloads than having no ratings. Only the lowest average ratings category were no different in number of downloads relative to no ratings at all; and contrary to what one might have expected, resources with low average ratings were not less likely to be downloaded than resources with no ratings. All the higher average ratings categories incrementally predicted significant increases in number of downloads: Hypothesis testing confirmed the significant difference between each of these categorizations ($p<.001$ for the difference between all categories except for the difference between average rating of 3-3.99 and 4-4.99 which was $p<.05$).

The number of ratings was also found to be a sizeable and statistically significant predictor of the variance. Number of ratings *per se* is an interesting factor because it is not *prima facie* an indicator of quality because ratings could be high or low. From the user perspective, more ratings could reflect greater interest in users (i.e., willing to rate) or simply social presence (others thought it worth downloading).

4. GENERAL DISCUSSION

4.1 The Individual Predictors

Similar to our previous work with TeacherPayTeachers [23], we found that the number of ratings and number of comments are separate predictors of resource downloads. This implies that a high number of low ratings predict more downloads than a resource with a low number of high ratings. This seemingly runs counterintuitive to the idea that more low ratings would dissuade teachers from looking at a resource. We are unsure as to the reason behind this finding and are currently pursuing additional research to further unpack this phenomenon.

We were also surprised to find that the current status of a resource's author would predict downloads but the length of the author's experience with TFA. This could be an indication of current TFA members being able to better understand other current members' resource needs. This could also be an indication of a continual change in teacher resource needs. For example, if a school administration changes curriculum then current teachers could have different resource needs than past teachers. However, if this hypothesis were true then we would have expected the date of a resource's upload to have a negative prediction on downloads. Instead, a resource's upload date was not a predictor.

The blue-ribbon status of a resource description is a reliable indicator of a quality resource since the designation is given by trained TFA administration. Because teachers are actively seeking only the best resources, it seems logical that blue ribbon status was found to be a significant predictor of future downloads.

File resources that are more easily editable meet an important need for teachers, mainly the ability to alter the resource for their specific instructional purposes. Consequently, file ease of editing was a significant predictor of future downloads.

Finally, the length of resource description was found to be a significant predictor of downloads. We suggest that this was the case since an author that writes a long description of their resource is probably the same type of author who provides ample detail in their resource proper.

4.2 The overall pattern of predictors

Based on our findings reported in this study, there is little to suggest that most TFA users rely on a same simple heuristic using *evaluative metadata* for choosing a resource to download. If this were the case, we would have expected to find one variable that predicted much more of the relative variance in downloads currently predicted by our model. As a whole, TFA teachers were influenced by many factors, seemingly following a weighted multiple regression decision making pattern.

However, from this aggregate level of analysis, we cannot conclude that the majority of users use complex decision heuristics because there remains the possibility that the heuristic relies on some other type of data. The predictive quality of the number of ratings and average ratings does cause us to hypothesize that TFANet users are likely relying on a complex heuristic to select a resource.

We hypothesize that the metadata that we identified as predicting download variance does so because it is used for individuals' stopping rule heuristics. This cannot be confirmed without further investigation into the motivations of online resource exchange users. Most obvious is the creation of a multilevel linear model that looks at the dependent variable of downloads as a growth model. This will be challenging given that resources are uploaded at different times. Yet this type of analysis, along with this paper and our previous work should eventually lead to a better understanding of the educative impacts of teacher resource exchanges.

4.3 Implications

We believe the findings in this paper could have two immediate impacts. Designers of resource exchanges, armed with the knowledge of how ratings predict downloads, can emphasize or de-emphasize resource characteristics in order to achieve desired behaviors. For instance, if a certain resource is not being downloaded then a designer can promote reviews for the resource which in turn might promote more interest and downloads.

These findings can also be used to generate educator professional development that better prepares resource exchange users to understand the role *evaluative metadata* plays in selecting resources. Developers of knowledge building communities [24] or of resource-based learning environments [8] can use these findings to better educate their participants into how to properly understand the role of *evaluative metadata*. One can imagine a resource exchange community using professional development based on this paper to help identify areas for improvement in data creation and review.

5. ACKNOWLEDGMENTS
Our thanks to Teach For America for generously providing us the data for our analysis.

6. REFERENCES
[1] Russell, M., Bebell, D., O'Dwyer, L. and O'Connor, K. Examining teacher technology use implications for preservice and inservice teacher preparation. *Journal of Teacher Education*, 54, 4 2003), 297-310.
[2] Recker, M., Dorward, J. and Nelson, L. Discovery and Use of Online Learning Resources: Case Study Findings. *Educational Technology & Society*, 7, 2 2004), 93-104.
[3] Poston, R. S. and Speier, C. Effective use of knowledge management systems: A process model of content ratings and credibility indicators. *MIS Quarterly*2005), 221-244.
[4] Mudambi, S. M. and Schuff, D. What makes a helpful online review? A study of customer reviews on Amazon. com. *MIS Quarterly*, 34, 1 2010), 185-200.
[5] Ball, D. and Cohen, D. *Developing practice, developing practitioners: Toward a practice-based theory of professional education*. Jossey-Bass Inc., City, 1999.
[6] Sparks, D. and Loucks-Horsley, S. Five models of staff development for teachers. *Journal of Staff Development*, 10, 4 1989), 40-57.
[7] Leat, D. and Higgins, S. The role of powerful pedagogical strategies in curriculum development. *Curriculum Journal*, 13, 1 2002), 71 - 85.
[8] Hill, J. and Hannafin, M. Teaching and learning in digital environments: The resurgence of resource-based learning. *Educational Technology Research and Development*, 49, 3 2001), 37-52.
[9] Schlager, M. S., Fusco, J. and Schank, P. *Evolution of an online education community of practice*. Cambridge University Press, City, 2002.
[10] Johnson, C. M. A survey of current research on online communities of practice. *The internet and higher education*, 4, 1 2001), 45-60.
[11] Morris, A. K. and Hiebert, J. Creating Shared Instructional Products. *Educational Researcher*, 40, 1 2011), 5.
[12] Sotiriou, S., Manouselis, N., Tzikopoulos, A., Costopoulou, C. and Sideridis, A. *Towards teacher competence on metadata and online resources: the case of agricultural learning resources*. Wageningen Academic Publishers, City, 2009.
[13] Schrum, L. Technology professional development for teachers. *Educational Technology Research and Development*, 47, 4 1999), 83-90.
[14] Ochoa, X. and Duval, E. *Use of contextualized attention metadata for ranking and recommending learning objects*. ACM, City, 2006.
[15] Ochoa, X. and Duval, E. Relevance ranking metrics for learning objects. *IEEE Transactions on Learning Technologies,* , 1, 1 2008), 34-48.
[16] Stein, M. K., Grover, B. W. and Henningsen, M. Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American educational research journal*, 33, 2 1996), 455-488.
[17] Davis, E. A. and Krajcik, J. S. Designing educative curriculum materials to promote teacher learning. *Educational Researcher*, 34, 3 2005), 3-14.
[18] Raab, M. and Gigerenzer, G. *Intelligence as smart heuristics*. Cambridge University Press, City, 2005.
[19] Todd, P. M. and Gigerenzer, G. Précis of simple heuristics that make us smart. *Behavioral and Brain Sciences*, 23, 5 2000), 727-741.
[20] Siegler, R. S. Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General*, 117, 3 1988), 258.
[21] Raudenbush, S. W. and Bryk, A. S. *Hierarchical Linear Models*. Sage Publications, Thousand Oaks, CA, 2002.
[22] Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T. and du Toit, M. *HLM 7*. Scientific Software International, Inc., City, 2011.
[23] Abramovich, S. and Schunn, C. Studying teacher selection of resources in an ultra-large scale interactive system: Does metadata guide the way? *Computers & Education*, 58, 1 2012), 551-559.
[24] Scardamalia, M. and Bereiter, C. Computer support for knowledge-building communities. *Journal of the Learning Sciences*, 3, 3 1994), 265-283.