

---

# **AC 2012-4126: AN INVESTIGATION OF DATA DISPLAYS FOR INTERPRETING PARTICIPATION IN ONLINE DISCUSSION: TWO PERSPECTIVES**

## **Erin Shaw, University of Southern California**

Erin Shaw is a Computer Scientist at the Information Sciences Institute at the University of Southern California's Viterbi School of Engineering. Her research focuses on modeling and assessing student knowledge in the areas of science and mathematics, and experimenting with new technologies for aiding assessment in distance learning. As a Co-principal Investigator on National Science Foundation-sponsored studies, she researches new ways to assess student collaboration in undergraduate engineering courses and new ways to motivate secondary mathematics learning in the context of computer game-making. Shaw was formally a Software Engineer in the field of computer graphics and taught math and science as a Peace Corps volunteer in Nepal. She has a bachelor's degree in mathematics from Fitchburg State University and a master's degree in computer graphics from Cornell University.

## **Dr. Micheal Crowley, University of Southern California**

## **Dr. Jaebong Yoo, University of Southern California**

He received a Ph.D. at Hallym University in Korea, 2009. He currently works as a Postdoctoral Researcher at Information Sciences Institute of the University of Southern California. His interests are educational data mining, intelligent transportation system, and multi-agent systems.

## **Hao Xu, University of Southern California**

Information Sciences Institute

## **Dr. Jihie Kim, University of Southern California**

Jihie Kim is the Principal Investigator of the Intelligent Technologies for Teaching and Learning group in the USC Information Sciences Institute (<http://ai.isi.edu/pedtek>). She is also a Research Assistant Professor in the Computer Science Department at the University Of Southern California (USC). Kim received a Ph.D. from the USC and master's and bachelor's degrees from the Seoul National University. Her current interests include pedagogical discourse analysis, human-computer interaction, social network assistance, and assessment of student collaborative online activities.

# An Investigation of Data Displays for Interpreting Participation in Online Discussion: Two Perspectives

## Abstract

This study investigated several types of data displays to determine which ones most effectively communicated information about participation in online discussions. The hypothesis was that improved data display would increase instructor efficiency with respect to the formative assessment of online student activities. Three types of assessments were examined in detail using data from undergraduate course discussion forums. The assessments displays were studied from two perspectives, that of the researchers and that of the instructor. Significant disparities were found between the assumptions of the researchers and the needs of the instructor. This work was part of a larger project to design e-learning workflows and reporting tools to monitor and interpret online course discussions for the purpose of instructional assessment.

## Introduction

The goal of the Pedagogical Workflows project is to develop an application that instructors can use to continuously monitor and assess online student discourse within their course discussion forums. The resulting Pedagogical Assessment Workflow System (PAWS) is a scalable e-learning framework that supports efficient and robust integration of diverse datasets for the purposes of student assessment. Datasets include discussion corpora, participation data, traditional assessment scores, survey responses, and demographic information from the course registrar. PAWS employs the same computational workflow technologies that support e-science applications in the fields of seismology and astronomy<sup>1</sup>. These existing workflow generation and execution approaches were applied to make online assessment accessible to instructors. PAWS' e-learning workflows were designed to produce results that answer assessment questions relevant to student discussions and provide formative feedback to instructors to facilitate "just in time" instructional adaptation to students learning and needs. The system is described in detail in previous papers<sup>2,3</sup>.

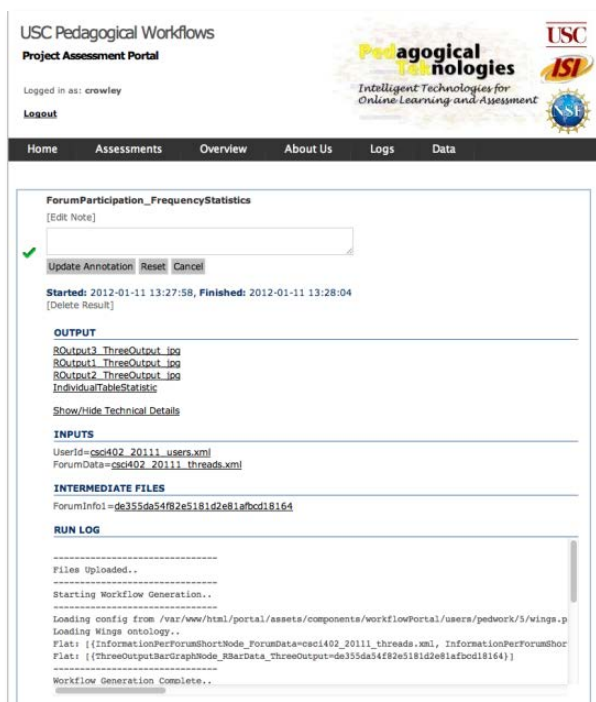


Figure 1. PAWS portal results page.

PAWS produces a results page that contains links to the data file and the table-based and graph-based output results, as well as logging information and an annotation facility. An example of a results page is shown in Figure 1. In the second year of the project, a Workflow Reporting and Feedback System (WRFS) was designed to automatically run instructor-selected workflows and deliver weekly reports via email<sup>4</sup>. The development of this system was motivated by the instructors' high interest in, but low use of the original workflow portal. The reports were linked to online forms through which teachers could respond to questions about the results and submit their feedback online. To determine if the results, especially the graphical representations, were optimal for instructional use, the team worked together with the instructor after each assessment, iteratively improving the data displays. The process and its outcomes are described in this paper. The study shows how greatly the perspectives of the research team and instructor differed initially, and how the differences were uncovered through repeated interviews with the instructor.

## Methodology

The workflow results examined in this paper were based on student discussions in an upper level undergraduate Operating Systems course at the University of Southern California. The course is a required course in the Computer Science curriculum and taught each semester. One of the paper's authors, Dr. Crowley, has taught the course since 1999. Between 80 and 150 students typically enroll each semester. The course requires considerable programming skills and time. Students wait in long lines during office hours and post hundreds of messages to the course discussion board each semester. With so many students requiring attention it is difficult for an instructor to attend to each individual's needs. The project team's goal was to create displays of the discussion data, especially graphical displays, to assist the instructor in assessing students' discussion forum participation at any given point in time. This study used data from the spring of 2011. The number of students participating in course discussions was 74.

This case study used a mixed method approach that included quantitative student participation statistics and qualitative feedback from interviews with the instructor. First, frequency data from an online discussion board was obtained. The workflow system can retrieve discussions from Blackboard, Moodle and phpBB discussion boards and this particular course used Moodle forums. The data corpus was updated nightly so that all displays were current as well as authentic. Descriptive statistics regarding the frequency of initial posts and responses were combined with information about the message author and the time waited for a response. The data was processed within PAWS, and the resulting output page included graphs and tables as well as processing information. The graphs were generated programmatically by the R statistics application.

Grounded theory style interviews were then conducted: In these, the course instructor was shown the results and asked the following three questions about each graph:

1. What does this graph tell you? Is it easy (difficult) to interpret?
2. How might you use this information to monitor and/or assess student activities?
3. How can the display and/or results be changed to make them more meaningful?

Questions were rephrased as necessary and follow up questions were asked whenever possible to clarify, and especially to explore ideas beyond the results we presented, to understand how we

might develop new and better results. The interview design was based on the modern view of grounded theory that emphasizes understanding through interpretation<sup>5</sup>. Creswell<sup>6</sup> described the approach as social constructivism, where the research intent is on making sense of meanings that others have about the world. It is an appropriate paradigm for studying instructors' perspectives of assessment and their interpretation of assessment results because of the strong relationships between the instructors and their unique instructional contexts. Notes from an initial interview were analyzed and new graphs were proposed. In a second analysis, investigators compared the new graph options and decided on a new suite of data displays. This suite of hand drawn graphs was then taken back to the instructor for a second round of interviews. The initial and resulting graphs are shown in the results sections.

## Results

In this section, we describe three types of analysis relating to discussion participation, the corresponding data displays, and how the displays evolved over the course of the interviews.

### Wait Time Analysis

“Wait Time” is how long students wait for responses to the questions they post. We know students become frustrated when questions are not answered in a reasonable amount of time, or when they are not answered satisfactorily. Knowing how long students wait might be an important factor when correlating discussion board use with performance and retention. As evaluators, we were interested in summative use, and in seeing a distribution based on wait time. We hypothesized that wait time might affect student interest, ability to complete projects, and retention in the course. The initial graph is shown in Figure 2.

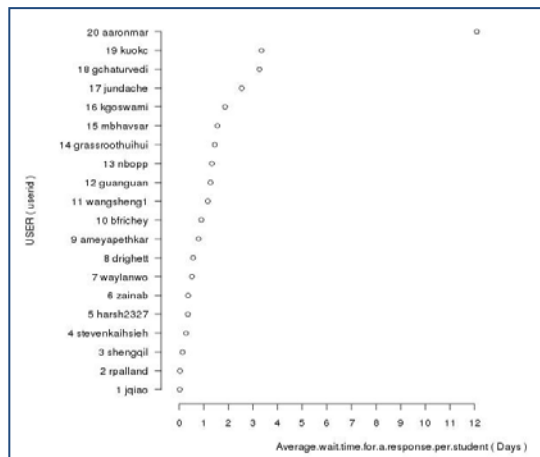


Figure 2. Original wait time analysis graph.

When shown the graph, the instructor commented that the scale was too big and that anything over three days was too old. The instructor wanted to see absolute wait time, not average time, which is biased by current time, and to know the thread subject for each open question, and have each question catalogued by forum. Importantly, he needed access to the thread itself so he could determine if the wait time was valid; for example, if the post was an assertion as opposed to an actual question. He also thought that showing questions by topic would be useful, as well as having a distribution of responses by topic over time (e.g. virtual memory was discussed over six hours.) Related to this, he suggested showing response curves, or thresholds, for answered questions by instructor, assistants and graders, to help him monitor how his course assistants helped answer student questions.

The final graph of this same data is actually a table, and is shown in Figure 3. It is partitioned into three sections, as shown: *Wait time until current time (unanswered questions)*, *Average wait time for a response (by user)*; and *Average wait time for a response (by forum)(not shown)*. Details such as links to messages were added so that the instructor could check the post. Note the

times shown in the tables, e.g. 356 days, are from re-running the workflow for this paper on last year's data; i.e., the numbers are always authentic. These were in fact, messages that were not answered, and the table makes the data clear.

USERID	WAIT TIME until Now(Unanswered)	Subject
swatigup	356 Day 20 Hours	<a href="#">Looking for a group</a>
kuock	356 Day 16 Hours	<a href="#">Experienced Computer Engineer looking to join a team</a>
amigoster	356 Day 4 Hours	<a href="#">Local Den CS Grad Student - Looking to be third member</a>
swatigup	354 Day 22 Hours	<a href="#">Looking to be a member of a group</a>
spchhabr	347 Day 21 Hours	<a href="#">Looking for a third partner.</a>
shopkins	339 Day 19 Hours	<a href="#">Output guideline typos</a>
supriya	336 Day 14 Hours	<a href="#">looking for a new team</a>
lokesh02	318 Day 17 Hours	<a href="#">Exception error which 5 type 0</a>
ameyapethkar@yahoo.co.in	298 Day 22 Hours	<a href="#">dirty bit</a>
metalsriks	269 Day 5 Hours	<a href="#">need of a registration server</a>
simon.zx	263 Day 2 Hours	<a href="#">Questions about Project 4</a>

USERID (Number of messages)	AVG WAIT TIME FOR A RESPONSE	Subject
aaronmar (1)	10 Day 2 Hours	Re: What's the best part about 402 projects?
kuock (3)	3 Day 8 Hours	Re: Question about Senator Behaviors Re: Name of file name for Exec() Re: Intel Career Events
gchaturvedi (1)	3 Day 6 Hours	Re: Remote Student Looking for a Group
mbhavsar (1)	1 Day 13 Hours	Re: Looking for Group partners
wangsheng1 (2)	1 Day 4 Hours	Re: looking for a group Re: looking for a third member for the team
kgoswami (3)	1 Day 2 Hours	Re: Regarding using Lock pointers Re: Question regarding FORK syscall Re: Swap File Clearing
bfrichey (1)	21 Hours	Re: Privileged vs. Regular Line
nbopp (2)	20 Hours	Re: Multiple Copies of Same Page in Memory Re: Total Ordering - Which

Figure 3. Final wait time analysis tables.

### Forum Participation Frequency Analysis

“Participation Frequency” is how frequently students use the discussion board, as well as how they use it. This data was shown over three bar charts, as shown in Figure 4: *Average number of messages per forum*, *Total number of messages per forum* and *Number of students participating in each forum*. As education researchers, our aim is to correlate these numbers to grades, and constructs such as motivation, and self-efficacy, and to study if participation affects learning outcomes. Each graph shows bars that correspond to numbers of initial posts, responses and total posts. The instructor suggested that longer threads might mean more student confusion, and generally speaking, more activity means more student problems. He commented that he didn't need to see all the forums, only the one theoretical (lecture) forum and the four project (assignment) forums were important. The assignments were mutually exclusive, so he was asking to be shown only two forums at any given time. This has implications for workflow processing, if individual forums must be specified.

The instructor also thought that having the three distinct bars was confusing, that the second legend wasn't necessary, and that the total post bar was redundant and could be omitted. He thought the bars were not clear without corresponding sample sizes, and generally thought the level was too high and that these graphs were not that useful for instructional assessment. For example, that knowing how many students participating in each forum is too high a level to be able to infer much. He indicated that data from each individual graph could be combined to show general activity level, and even that the same axis could be used. Related to this, he suggested using distribution rather than average, for example, using bars to represent the number of messages in each thread, or showing a group of bars, where bar height was length (i.e., number of messages) for each topic discussed. The final graph of this data is shown in Figure 5.

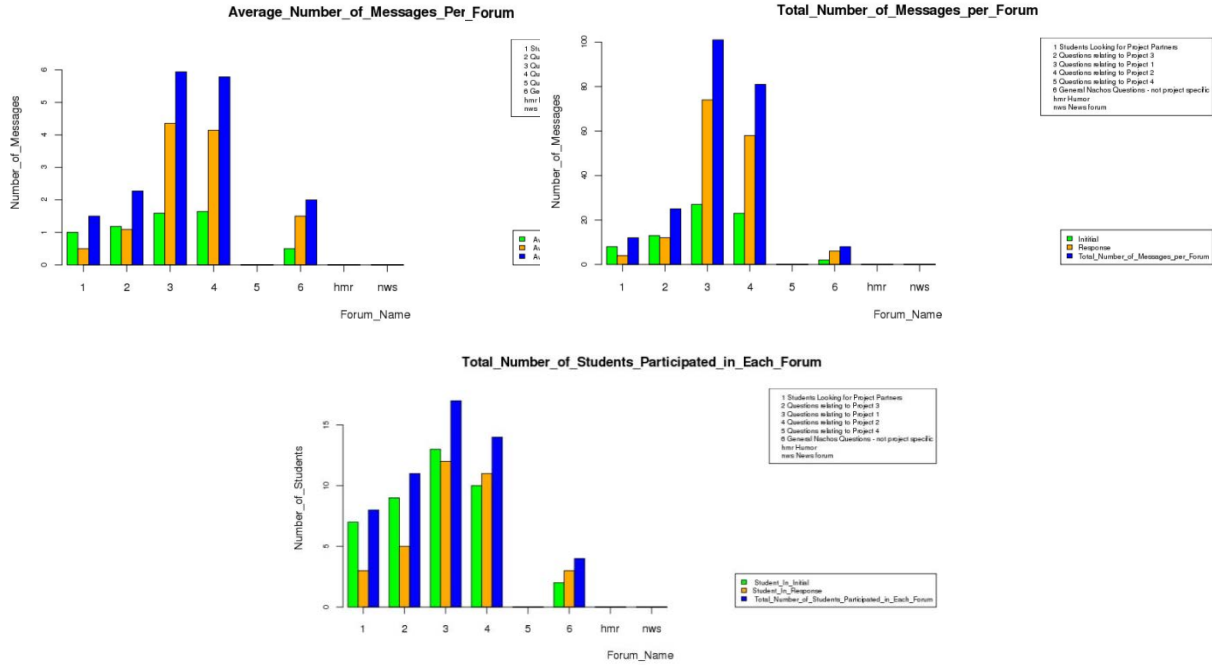


Figure 4. Original three forum participation analysis graphs.

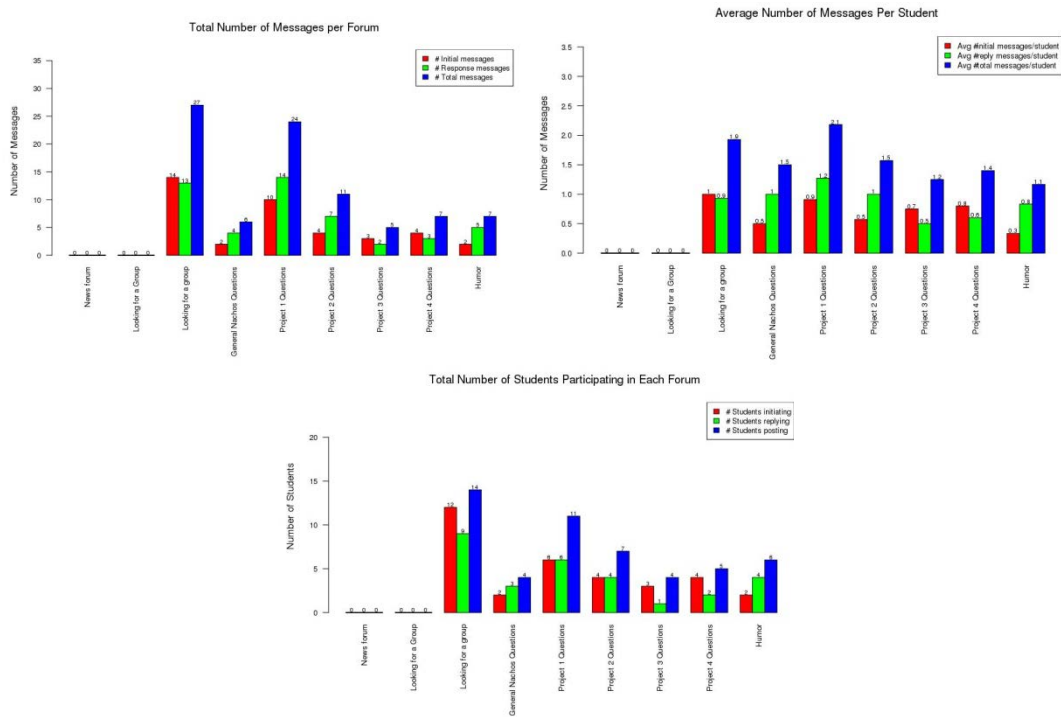
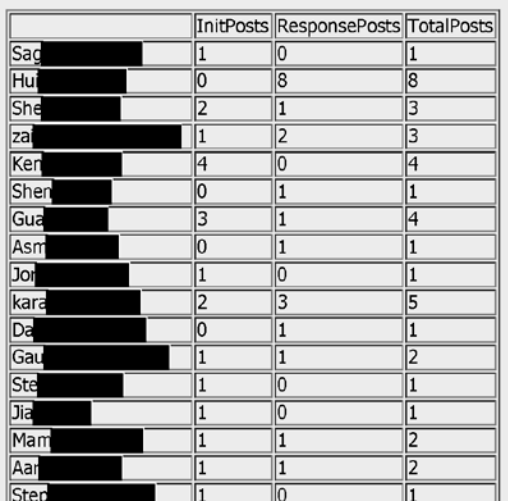


Figure 5. Final three forum participation analysis graphs.

## Student Participation Analysis

“Student Participation” is the traditional, per student, descriptive statistic that is commonly displayed by course management systems. The workflow system outputs information about student participation as a table, as shown in Figure 7. Like participation frequency statistics, these numbers are typically correlated with grades and constructs such as motivation, and self-efficacy. The instructor did not feel these statistics were that useful, but suggested that they could be used to monitor graders (one of the names listed was inadvertently that of a grader and not a student). He also thought that *question and answer* statistics (as opposed to *initial post and response post* statistics, i.e., the differentiation of the content as opposed to the position of the post) would be useful. Content differentiation was recently added to the workflow suite using a classifier based on machine learning<sup>13</sup>.



	InitPosts	ResponsePosts	TotalPosts
Sac	1	0	1
Hu	0	8	8
She	2	1	3
za	1	2	3
Ker	4	0	4
Shen	0	1	1
Gua	3	1	4
Asm	0	1	1
Jon	1	0	1
kara	2	3	5
Da	0	1	1
Gau	1	1	2
Ste	1	0	1
Jia	1	0	1
Mam	1	1	2
Aar	1	1	2
Ste	1	0	1

Figure 7. Student participation frequency.

The instructor also suggested that, analogous to showing initial posts and response posts *BY* the student, we might show response posts *TO* the students’ posts, and, furthermore, break these down by role, i.e. whether the response was from a student, assistant, or instructor. Related to this, the instructor suggested that we show the thread length for response (or answer). He also wanted to know, if an answer was good, why were there more answers? This table has not yet been updated to reflect the results of the investigation.

## **Scholarly Significance**

There is an emphasis today on data-driven instruction; the data referred to typically ranges from summative, standardized exam data, to weekly quiz results that are used for formative assessment purposes. However, participation in partially or fully online education courses continues to grow, and there is a growing need for in-depth assessment of online forum participation and discourse. We can draw from several studies to motivate the exploration of practitioner-driven assessment. In Black & Wiliam’s<sup>7</sup> foundational work on formative assessment, for example, they concluded “the effective development of formative assessment would come about only if ‘each teacher finds his or her own ways of incorporating the lessons and ideas ... into her or his own patterns of classroom work.’” (p.20). And as Conole et al.<sup>8</sup> observed, despite the plethora of theoretical frameworks that have been applied to e-learning there is “little evidence of how these models or theories are applied to effective pedagogically driven e-learning”. Delivering practitioner-friendly analytical results to practitioner-based assessment questions is the key to making formative assessment work; this work provides evidence for how this particular type of e-learning assessment can be designed to be most effective.

## **Related Work**

Graphs that are not easily understood by users are all too common. Goodman & Hambleton<sup>9</sup> remark on the phenomenon in the area of general student assessment, reporting that “a great amount of attention has been directed toward the creation of technically sound assessment” (p. 145) but that “considerably less attention” has been paid to organizing, reporting and using the results. They found that the use of graphical displays, among other less relevant (to this paper) characteristics appear to be effective in making reports more “concrete and meaningful” to readers. The finding was supported by Miller and Watkins<sup>10</sup>, where many of their examples of “good graphs” are simple bar graphs. Users in the study were parents, however, while in this study they are engineering instructors. An alternative to using asynchronous iterative design is described in Perer and Sneiderman’s<sup>11,12</sup> seminal work on guiding exploratory data analysis, in which they described architecture to support social and collaborative graph annotation to facilitate data understanding by social scientists and intelligence analysts. Each of these communities desired to understand data from their unique perspectives, and came to an understanding based on the underlying flexibility of the tools. However, sophisticated data analysis tools are expensive and instructional needs are unique to each class. Moreover, because of the instructor’s unique relationship with the course data, we felt it was important to use current and authentic data. We do, however, plan to interview different instructors about this same data in the future, and so will be looking at this issue in more depth.

## **Conclusions**

The Pedagogical Workflows team has been developing graphs to explain the results of participation assessment analysis. This study has illustrated the challenge of using graphs, and visualizations in general, to summarize data as part of an instructional reporting system. There may be multiple stakeholders and thus multiple interpretations of the results. It demonstrated the necessity of working collaboratively with stakeholders, of using interview techniques that facilitated understanding through dialog, feedback and analysis to develop graphical results that were meaningful to a different kind of user – an instructor, as opposed to an education researcher. The difference of regard with respect to the time scale and frequency representation (i.e., averaged versus absolute) was especially notable. Moreover, these results may only be meaningful to specific instructors, given the unique nature of any one course, although we expect that instructors who use question and answer style discussion boards will also find these results useful. The next step in the study is to interview a second teacher, whose course workflows have been developed, starting with the results from this investigation.

## **Acknowledgements**

The work was supported by the National Science Foundation, under Human-Centered Computing grant [#0917328](#).



## Bibliography

1. Deelman, E., Singh, G., Su, M., Blythe, J., Gil, Y., Kesselman, C., Mehta, G., Vahi, K., Berriman, G. B., Good, J., Laity, A., Jacob, J. C., and D. S. Katz (2005). Pegasus: a Framework for Mapping Complex Scientific Workflows onto Distributed Systems, *Scientific Programming Journal*, Vol 13(3), (2005).
2. Ma, J., Kang, J.-H., Shaw, E. and Kim, J. (2011), [Workflow-Based Assessment of Student Online Activities with Topic and Dialogue Role Classification](#), *Proceedings of the AI in Education Conference*, 2011.
3. Ma, J., Shaw, E. and Kim, J. (2010) [Computational Workflows for Assessing Student Learning](#), *The 10th International Conference on Intelligent Tutoring Systems Conference*, 2010.
4. Xu, H., Shaw, E. and Kim, J. (in process) Workflow reporting and feedback system for formative assessment of online course discussions.
5. Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. London: SAGE.
6. Creswell, J.W. (2007). *Qualitative Inquiry & Research Design: Choosing Among Five Approaches*, p.20-21, SAGE.
7. Black, P. & Wiliam, D. (1998). *Assessment in Education: Principles, Policy & Practice*, 1465-329X, Volume 5, Issue 1, 1998, Pages 7 – 74.
8. Conole, G., Dyke, M., Oliver, M. and Seale, J. (2004) 'Mapping pedagogy and tools for effective learning design', *Computers & Education*, vol. 43, nos. 1–2, pp. 17–33.
9. Goodman & Hambleton (2004). Student Test Score Reports and Interpretive Guides: Review of Current Practices and Suggestions for Future Research. *Applied Measurement in Education*, 17(2), 145-220.
10. Miller, J.A. and Watkins, M.W. (2010). The Use of Graphs to Communicate Psycho-educational Test Results to Parents. *Journal of Applied School Psychology*, Volume 26, Issue 1.
11. Perer, A. and Shneiderman, B. (2008a) Systematic Yet Flexible Discovery: Guiding Domain Experts through Exploratory Data Analysis. In *Proc. Intelligent User Interfaces (IUI)*. ACM Press (2008).
12. Perer, A. and Shneiderman, B. (2008b) Integrating Statistics and Visualization: Case Studies of Gaining Clarity during Exploratory Data Analysis. (Under Submission)(2007).
13. Soo Won Seo, Jeon-Hyung Kang, Joanna Drummond and Jihie Kim, Using Graphical Models to Classify Dialogue Transition in Online Q&A Discussions, In *Proceedings of the AI in Education Conference*, 2011.