From Data Science to Learning Analytics

Evolution of the E-Listening Research Project

DATA

THEORY

COMPUTATION

DISPLAY

USE

Dr. Alyssa Wise

Associate Professor

Simon Fraser University

April 13, 2016

DATA

THEORY

COMPUTATION

DISPLAY

USE

Detecting different ways learners attend to others' posts in online discussions and how they relate to the contributions they make themselves as an element of dialogic participation

DATA

THEORY

COMPUTATION

DISPLAY

USE

E-Listening Research Project

Online Listening Theory

DATA

THEORY

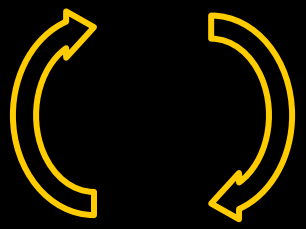
COMPUTATION

DISPLAY

USE

- Goal of online discussions is for learners to build understanding through dialoging with others
- Drawing on (neo-piagetian) theories of social constructivism at a basic level this involves

Externalizing one's ideas by contributing posts to an online discussion



Taking in the externalizations of others by accessing existing posts

Wise et al., 2012a; Wise et al., 2013

Listening Not Lurkers

DATA

THEORY

COMPUTATION

DISPLAY

USE

Lurker	Listening
A person	A behavior
w/ nefarious connotations	that is normal / productive,
who looks at posts as a stand alone activity	part of a larger engagement in discussion,
and doesn't contribute themselves	and can inform future discussion contributions

Wise et al., 2013; Muller et al., 2010; Preece et al., 2004

E-Listening
Research Project

Not "if" but "how"

listen

DATA

THEORY

COMPUTATION

DISPLAY

USE

- Prior research assumed students generally attend to others' posts
- Yet evidence of limited attention to previous discussion posts
- Our data-mining work shows that learners make widely varying choices in the decision space
  - Frequency /length of log-in
  - Which posts attended to, in what order, for how long
  - Revisit as often as wanted
  - Reply when + where desired

Wise et al., 2013; 2012a, Brooks et al., 2013; Palmer, et al. 2008; Dennen, 2008; Hewitt, 2003;

E-Listening Research Project

Collecting Listening Data

DATA

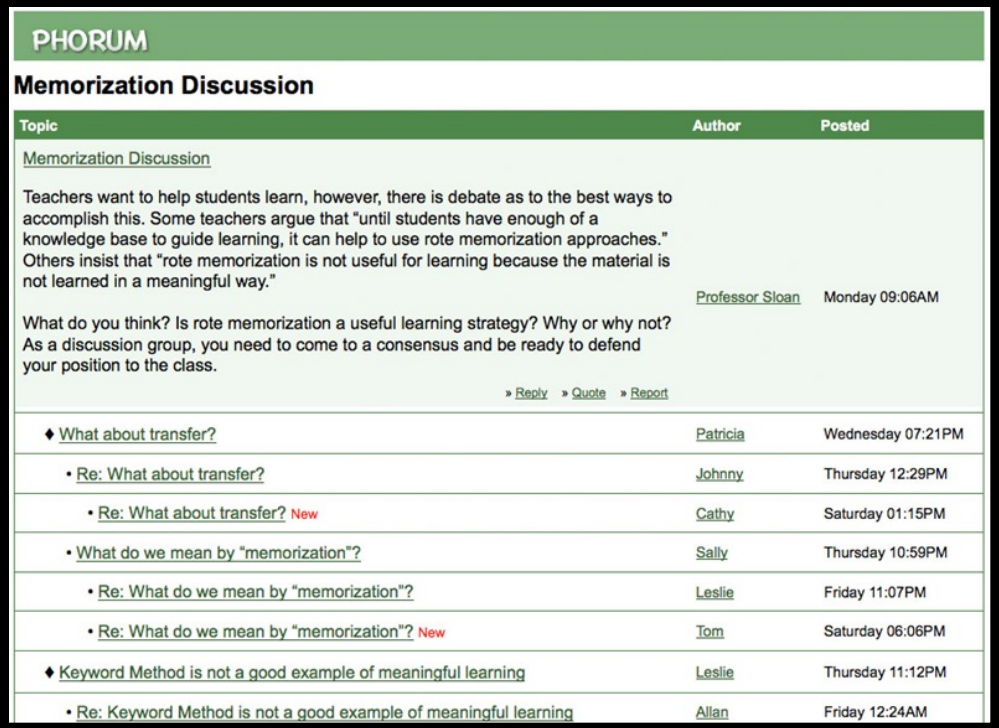
THEORY

COMPUTATION

DISPLAY

USE

- Online discussion forums don't naturally collect listening data
- Need to know what specific post a student is looking at
- Modifications made to open-source forum "Phorum"
- Collect ~1000 + records per student



PHORUM

Memorization Discussion

Topic	Author	Posted
Memorization Discussion Teachers want to help students learn, however, there is debate as to the best ways to accomplish this. Some teachers argue that "until students have enough of a knowledge base to guide learning, it can help to use rote memorization approaches." Others insist that "rote memorization is not useful for learning because the material is not learned in a meaningful way." What do you think? Is rote memorization a useful learning strategy? Why or why not? As a discussion group, you need to come to a consensus and be ready to defend your position to the class. » Reply » Quote » Report	Professor Sloan	Monday 09:06AM
♦ What about transfer?	Patricia	Wednesday 07:21PM
• Re: What about transfer?	Johnny	Thursday 12:29PM
• Re: What about transfer? New	Cathy	Saturday 01:15PM
• What do we mean by "memorization"?	Sally	Thursday 10:59PM
• Re: What do we mean by "memorization"?	Leslie	Friday 11:07PM
• Re: What do we mean by "memorization"? New	Tom	Saturday 06:06PM
♦ Keyword Method is not a good example of meaningful learning	Leslie	Thursday 11:12PM
• Re: Keyword Method is not a good example of meaningful learning	Allan	Friday 12:24AM

Data Processing Needs

- DATA
- THEORY
- COMPUTATION
- DISPLAY
- USE

- Turn comprehensive point-in-time entries into individual activity records
- Coding events and calculating duration
  - Clicks can initiate or complete actions
  - Conceptual event categories
    - Scan vs. read (6.5 wps threshold)
    - Post and review (self-oriented)
- Divide activity into sessions using abandonment criteria (time of inactivity)
- Determine time window for variable calculation (study-dependent – could be overall or based on learning activity, fixed time, individual student session)

Data Processing Techniques

DATA

THEORY

COMPUTATION

DISPLAY

USE

- mySQL queries merging log + post tables produces list of all actions, time-date stamp, ID of user performing action, post acted on, ID of user who created post
- Excel VBA macros clean data, separate users, calculate initial event duration, divide sessions of use, estimate session-ending events (avg. speed (user| event) x post length), recode reads and reviews
- Macros also calculate variables at desired unit of analysis –(e.g. # of sessions, av. session length, % sessions with posts, % other's posts viewed (or read), % real reads, av. real read length, # of reviews (self/other) , av. review length)

Wise & Chiu, 2014; Wise et al. 2014; 2013; 2012a

E-Listening Research Project

Modelling Techniques

- Cluster Analysis used to identify and characterize learner subpopulations
- HLM Mixed Modeling used to examine listening-speaking relationships + take into account potential interdependencies of students collaborating in groups
- Statistical Discourse Analysis used to detect distinct segments of discussion and build explanatory models of these

Wise & Hsiao, in preparation; Wise et al. 2014; 2013; Wise & Chiu, 2014; Wise & Chiu, 2011

DATA
THEORY
COMPUTATION
DISPLAY
USE

E-Listening Research Project

Studies + Findings

DATA

THEORY

COMPUTATION

DISPLAY

USE

Research conducted on large UG courses in business + ed psych (online + blended offerings)

Common Patterns	Characteristic Behaviors
Disregardful	Minimal attention to others' posts (few posts viewed; short time viewing). Brief and relatively infrequent sessions of activity.
Coverage	View a large proportion of others' posts, but spends little time attending to them (often only scanning). Short but frequent sessions of activity, focusing primarily on new posts. **May be socially-oriented or content-driven.*
Focused	View limited number of others' posts, but spends substantial time attending to them. Few extended sessions of activity.
Thorough	Views a large proportion of other's posts + spends substantial time attending to many of them. Long overall time spent listening; considerable reviews of others' posts.

Wise & Chiu, 2014; Wise et al. 2014; 2013; 2012a; 2012c;

E-Listening Research Project

Studies + Findings

DATA

THEORY

COMPUTATION

DISPLAY

USE

Research conducted on large UG courses in business + ed psych (online + blended offerings)

Dimension	Characteristic Behaviors
Breadth	The **quantity of unique posts** one views - important in terms of the **diversity of ideas** a learner is exposed to.
Depth	The **length of time** spent reading posts – important to allow for **deeper consideration** of others' ideas.
Temporal Contiguity	The degree to which **learners disperse or concentrate** their participation - important for integration / evolution over time.
Revisitation	The extent to which students **return to posts** made by themselves and others – important in metacognition and self-regulation.

Wise & Chiu, 2014; Wise et al. 2014; 2013; 2012a; 2012c;

E-Listening Research Project

Studies + Findings

DATA

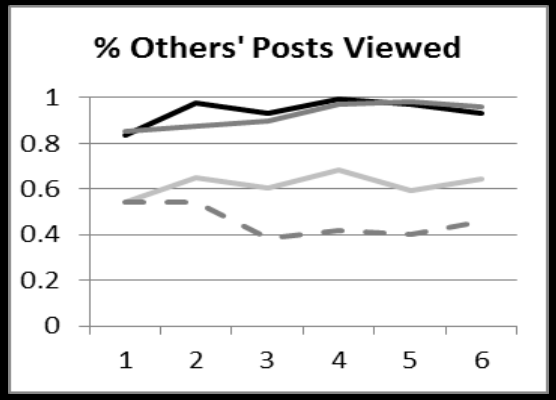
THEORY

COMPUTATION

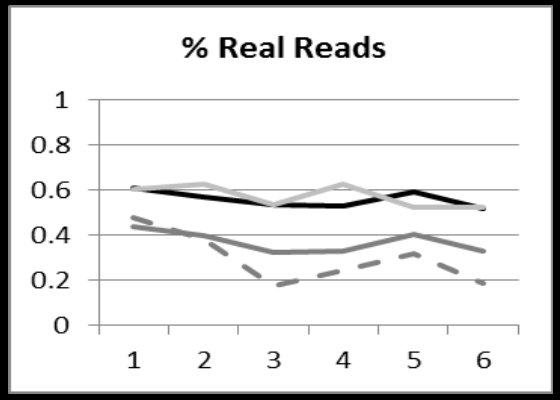
DISPLAY

USE

Breadth



Depth



Depth	Breadth: *Low*	Breadth: *High*
Low	**Disregardful**	**Coverage**
High	**Focused**	**Thorough**

Wise & Chiu, 2014; Wise et al. 2014; 2013; 2012a; 2012c;

E-Listening Research Project

Studies + Findings

DATA

THEORY

COMPUTATION

DISPLAY

USE

Relationships

Listening Depth

- A greater % of real reads predicts richer argumentation (reasoning/qualifiers)

Listening Reflectivity

- Reviewing others' posts multiple times predicts greater responsiveness
- Reviewing one's own posts multiple times predicts richer argumentation (# of claims made)

Listening Breadth

- Reading a greater % of posts and viewing a greater % of posts than those read predicts richer argumentation (evidence used to support claims)

Wise & Chiu, 2014; Wise et al. 2014; 2013; 2012a; 2012c;

Studies + Findings

DATA

THEORY

COMPUTATION

DISPLAY

USE

Interventions

Assigning Students Roles to Play

- Synthesizer / Wrapper roles increased the % of posts read during in-role weeks (and sometimes the % of real reads), but effect weakly sustained post-role
- Roles often contributed extensive summaries that acted as pivotal posts spurring whole discussion to more advanced knowledge construction phase

Open-Ended vs Distinct Alternative Tasks

- Open-ended business challenge led to a higher % of real reads and average time spent reviewing one's own posts
- Follow-on argumentation effects possible

Wise & Hsiao, in prep; Wise & Chiu, 2014; Wise & Chiu, 2011;

DATA

THEORY

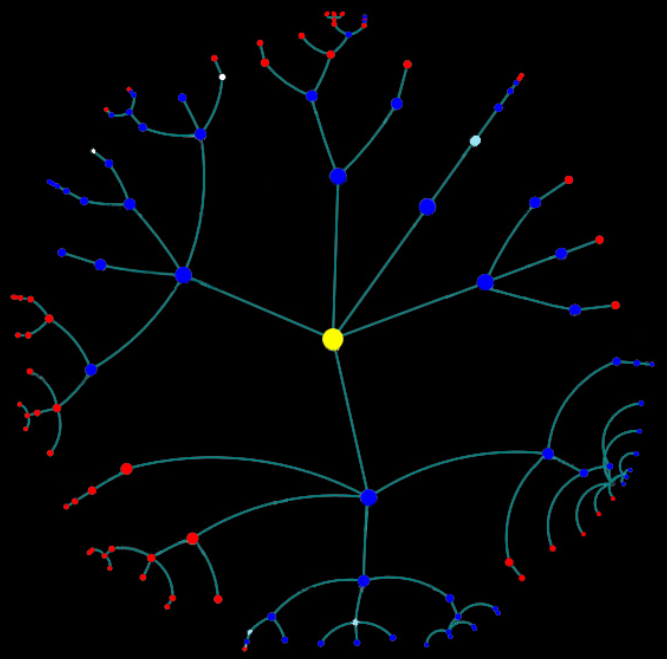
COMPUTATION

DISPLAY

USE

Starburst

A redesigned online discussion forum that includes embedded visual analytics of conversation structure + students' interactions



DATA

THEORY

COMPUTATION

DISPLAY

USE

Visualizing Discussion Analytics

DATA

THEORY

COMPUTATION

DISPLAY

USE

Embedded Analytics

Integration of activity traces into existing learning environment interface

[Stands in contrast to extracted analytic systems such as dashboard that set-up analytics use apart from other aspects of the learning activity]

Goal is real-time use of analytics to guide student activity + choices to support learning

The following example builds on some of the listening analytics presented earlier

Online Discussion Challenges

DATA

THEORY

COMPUTATION

DISPLAY

USE

- Students feel overwhelmed, have difficulties navigating + deciding which posts to read/reply to
- Often they only view new posts and reply to most recent ones
- Low levels of interactivity
- Fractured and incoherent conversations
- Unintentional discussion drift and thread death; discussions never converge to a conclusion

Wise et al., 2012b; Peters & Hewitt, 2010; Hewitt, 2005; Dringus & Ellis, 2005; Reyes & Tchounikine, 2003

Online Discussion Challenges

DATA

THEORY

COMPUTATION

DISPLAY

USE

Linear

Text-based

Thread Indenting

Hard to Scan

No Sense of Whole

Interaction Info not Salient

Visualizing Discussion Analytics

DATA

THEORY

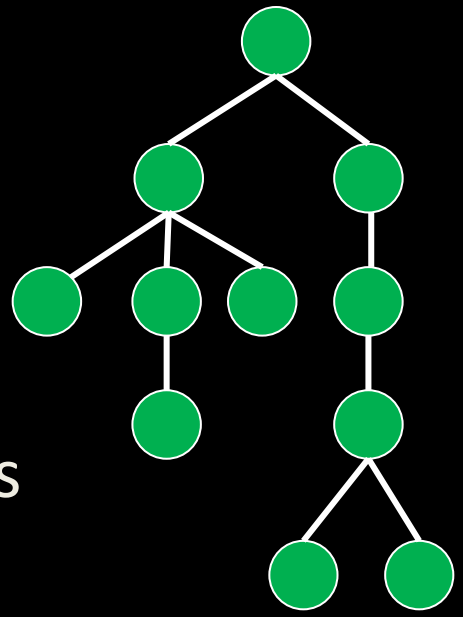
COMPUTATION

DISPLAY

USE

DG1: Represent discussion structure

- Match between discussion and "tree" structures
- Post = Node
- Reply = Link
- Acyclic network of connected nodes (no loops)



- Multiple children, one parent
- Used in discussion analysis

Wise & Padmanabhan, 2009; Teplovs, 2008; Scardamalia, 2004; Aviv, et al., 2003; Hara et al., 2000

Visualizing Discussion Analytics

DATA

THEORY

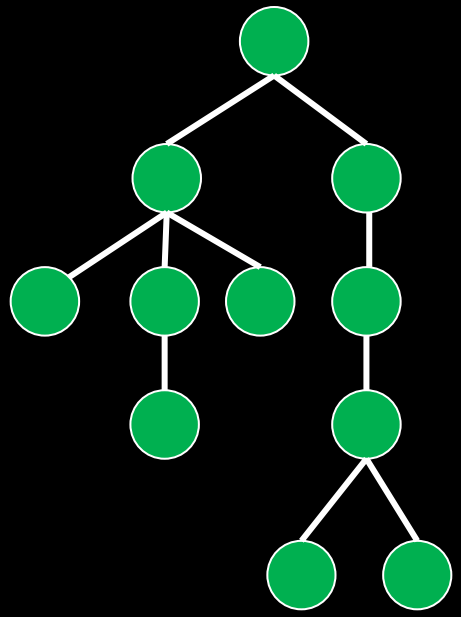
COMPUTATION

DISPLAY

USE

DG1: Represent discussion structure

- Match between discussion and "tree" structures



Visualizing Discussion Analytics

DATA

THEORY

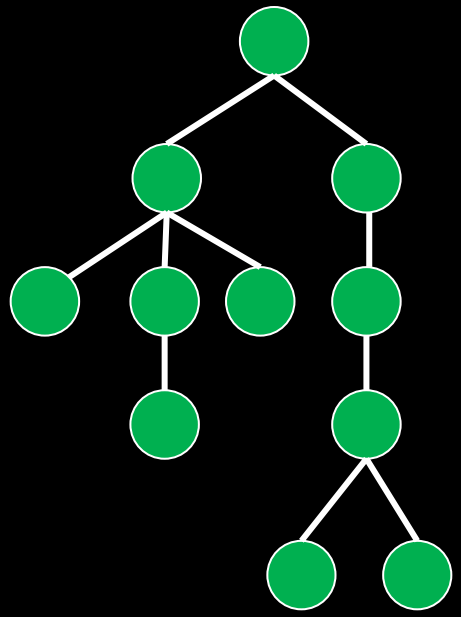
COMPUTATION

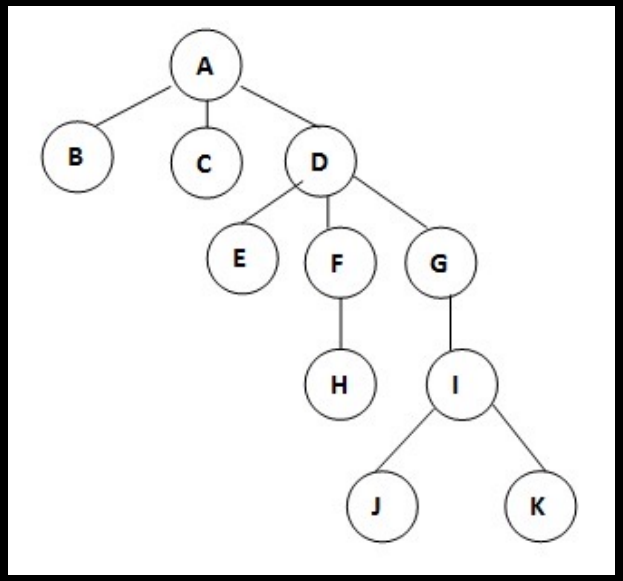
DISPLAY

USE

DG1: Represent discussion structure

- Match between discussion and "tree" structures





Visualizing Discussion Analytics

DATA

THEORY

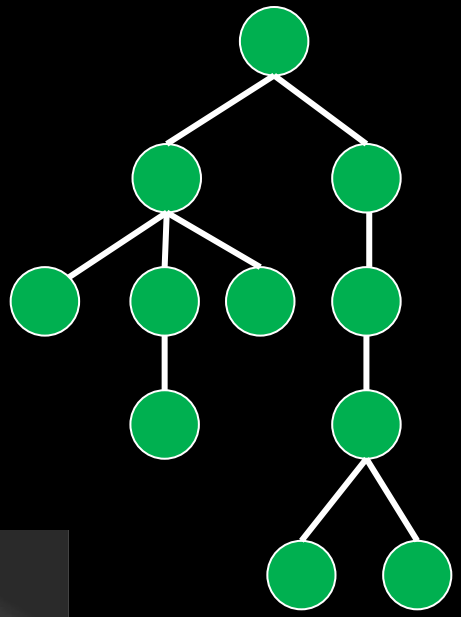
COMPUTATION

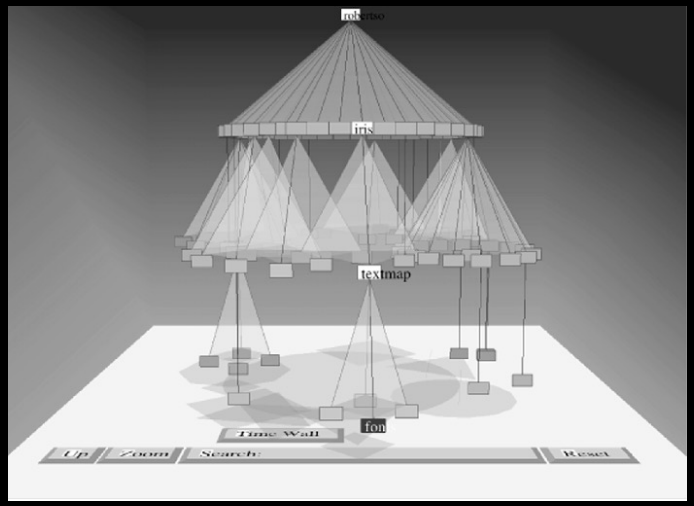
DISPLAY

USE

DG1: Represent discussion structure

- Match between discussion and "tree" structures





Visualizing Discussion Analytics

DATA

THEORY

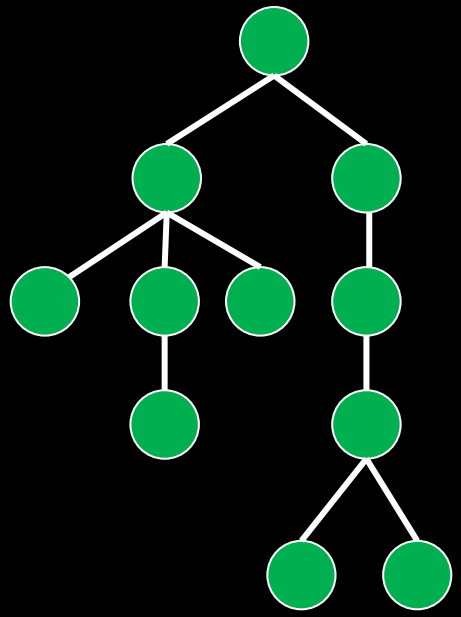
COMPUTATION

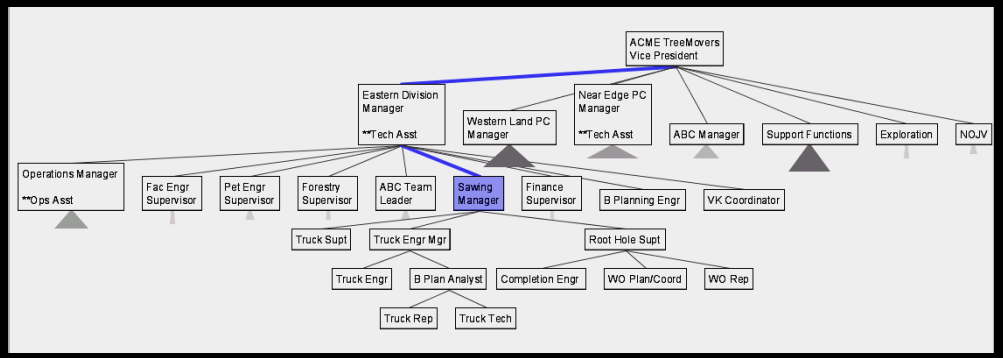
DISPLAY

USE

DG1: Represent discussion structure

- Match between discussion and "tree" structures





Visualizing Discussion Analytics

DATA

THEORY

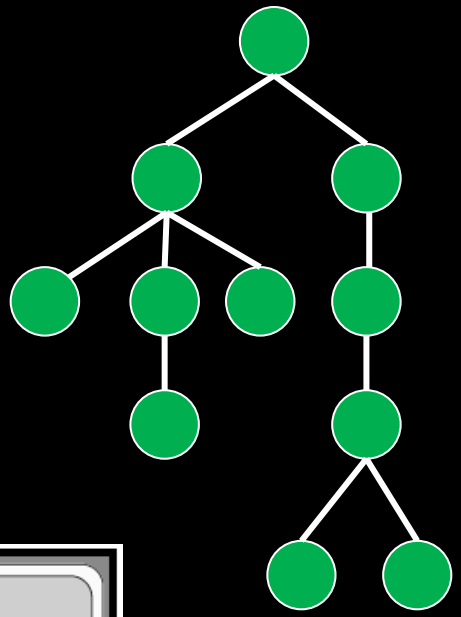
COMPUTATION

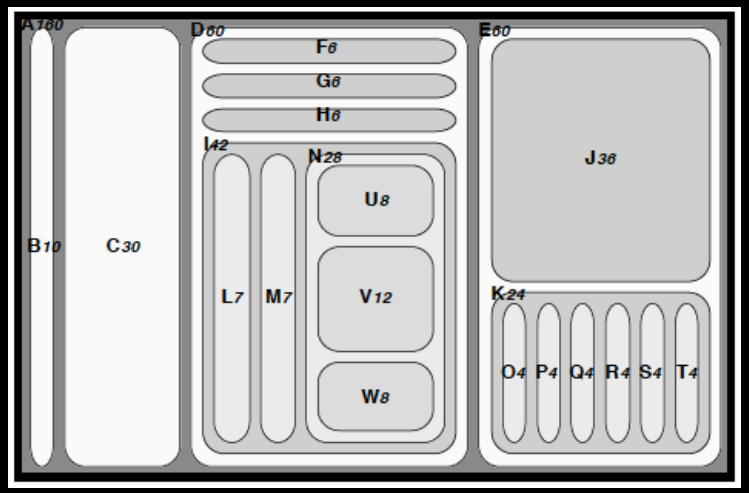
DISPLAY

USE

DG1: Represent discussion structure

- Match between discussion and "tree" structures





Visualizing Discussion Analytics

DATA

THEORY

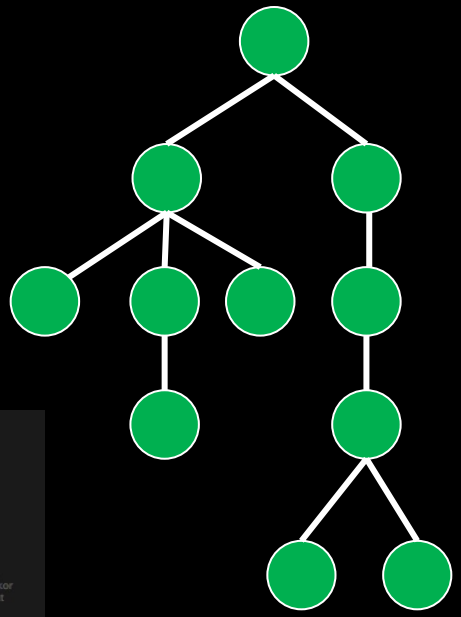
COMPUTATION

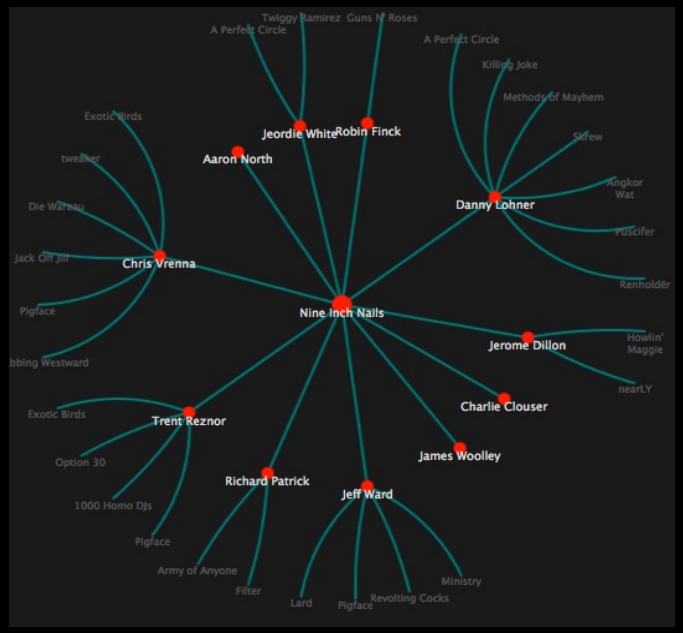
DISPLAY

USE

DG1: Represent discussion structure

- Match between discussion and “tree” structures





Visualizing Discussion Analytics

DATA

THEORY

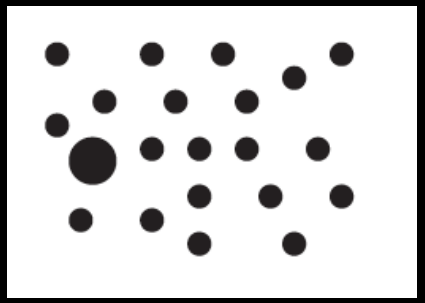
COMPUTATION

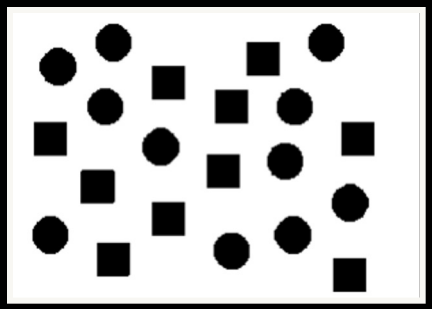
DISPLAY

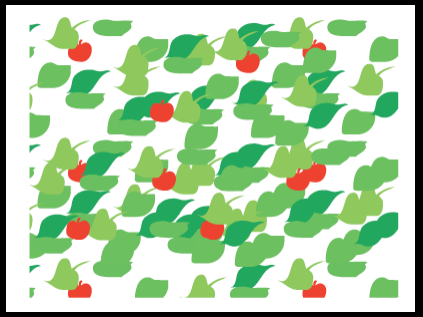
USE

DG2: Embed post access info

- Read / Viewed / Unopened
- Self / Peer / Instructor Post
- Subject and Date of Post







Visualizing Discussion Analytics

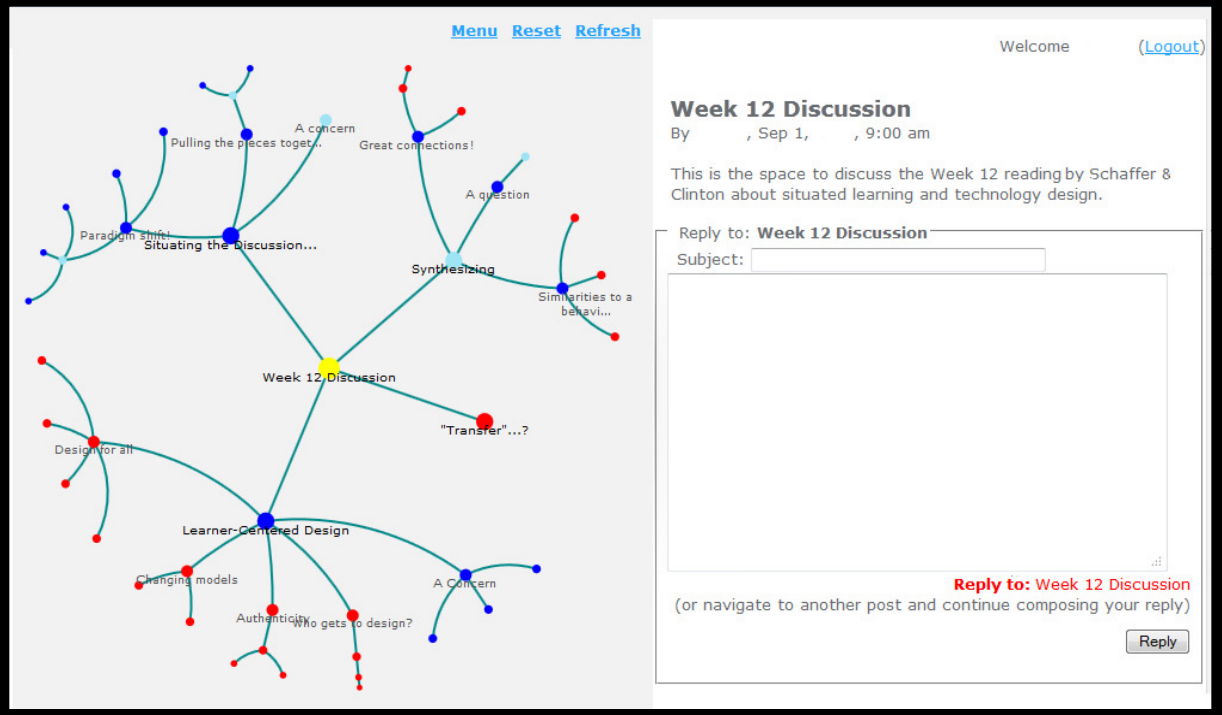
DATA

THEORY

COMPUTATION

DISPLAY

USE



Embedded Analytics Use

DATA

THEORY

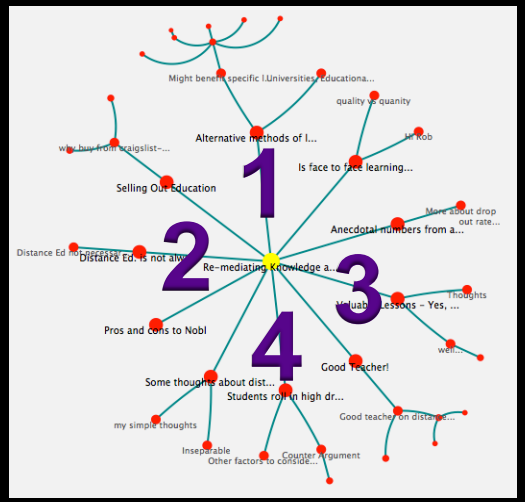
COMPUTATION

DISPLAY

USE

Test 1: Comparative case studies using clickstream data in lab setting

- Students interacted with same authentic discussion in Starburst and traditional interface
- With embedded analytics students were more active in selecting which threads to visit



Marbouti & Wise, 2016

Embedded Analytics Use

DATA

THEORY

COMPUTATION

DISPLAY

USE

Test 2: Live implementation in blended graduate seminar

- Students reported Starburst was helpful for seeing discussion flow and made it easy to get a sense overall + decide where to go

"I jumped around to get a sense of things, then dug in deep"

"As a facilitator you could see where attention was needed"

"Easy to see replies to my posts"

"I would decide to visited more isolated or populated threads depending on time"

Wise et al., 2014

DATA

THEORY

COMPUTATION

DISPLAY

USE

Aligned Design

Design-based exploration of the issues involved in how analytics are taken up and used by students as part of their process of learning

Aligned Design

DATA

THEORY

COMPUTATION

DISPLAY

USE

Aligned Design

Challenges in Analytics Use

DATA

THEORY

COMPUTATION

DISPLAY

USE

Challenges of Interpretation

- *Context*
- *Trust*
- *Priorities*
- *Individuality*

Challenges of Action

- *Possible Options*
- *Enacting Change*
- *Dependency*

Wise et al., in press; Wise, 2014

Aligned Design

Student Tuning Model

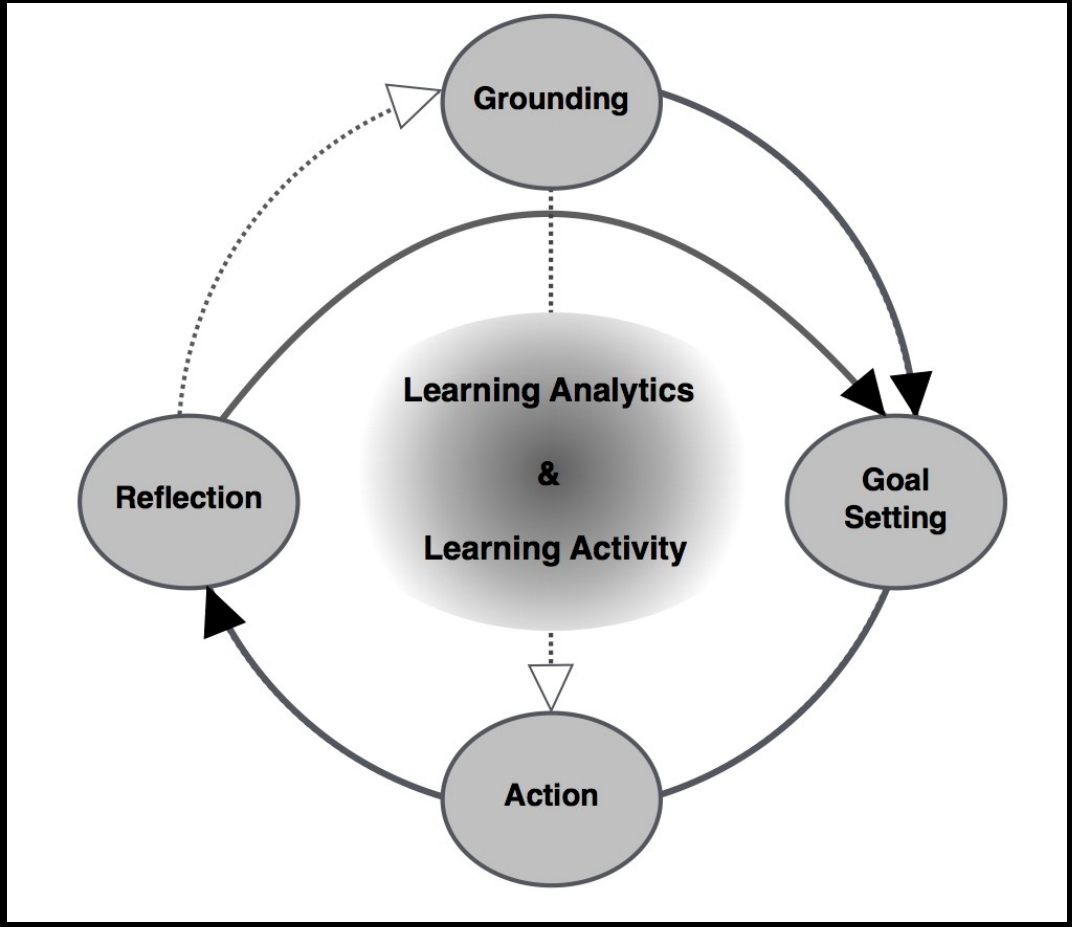
DATA

THEORY

COMPUTATION

DISPLAY

USE



A metacognitive cycle for analytics use drawing on the self-regulation and reflective practitioner literatures

Wise et al., in press, Wise, 2014

Aligned Design

Align Design Framework

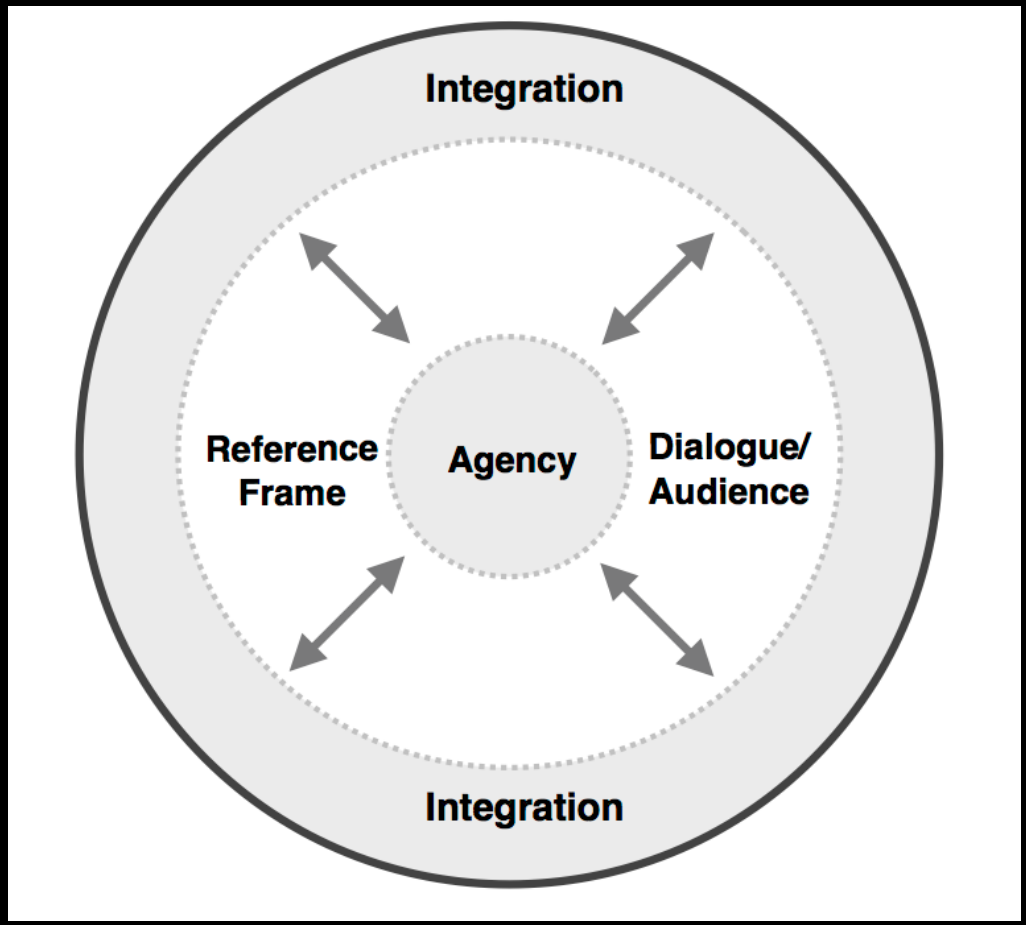
DATA

THEORY

COMPUTATION

DISPLAY

USE



A series of principles that raise sensitizing issues for the design of analytics use and opens a collective space for investigation

Wise et al., in press; Wise, 2014

INTEGRATION

Clear guidelines and discussion of

- the purpose of engaging in online discussion

  articulating one's ideas, being exposed to the ideas of others, negotiating differences in perspective

- the instructor's expectations for a productive process of engaging in online discussions

  attending deeply to a spectrum of others' ideas, and contributing comments that are responsive and rationaled,

- how the learning analytics provide indicators of this process

  percent of posts read introduced as a metric that has clear meaning in the context of the activity

INTEGRATION

Aligned Design

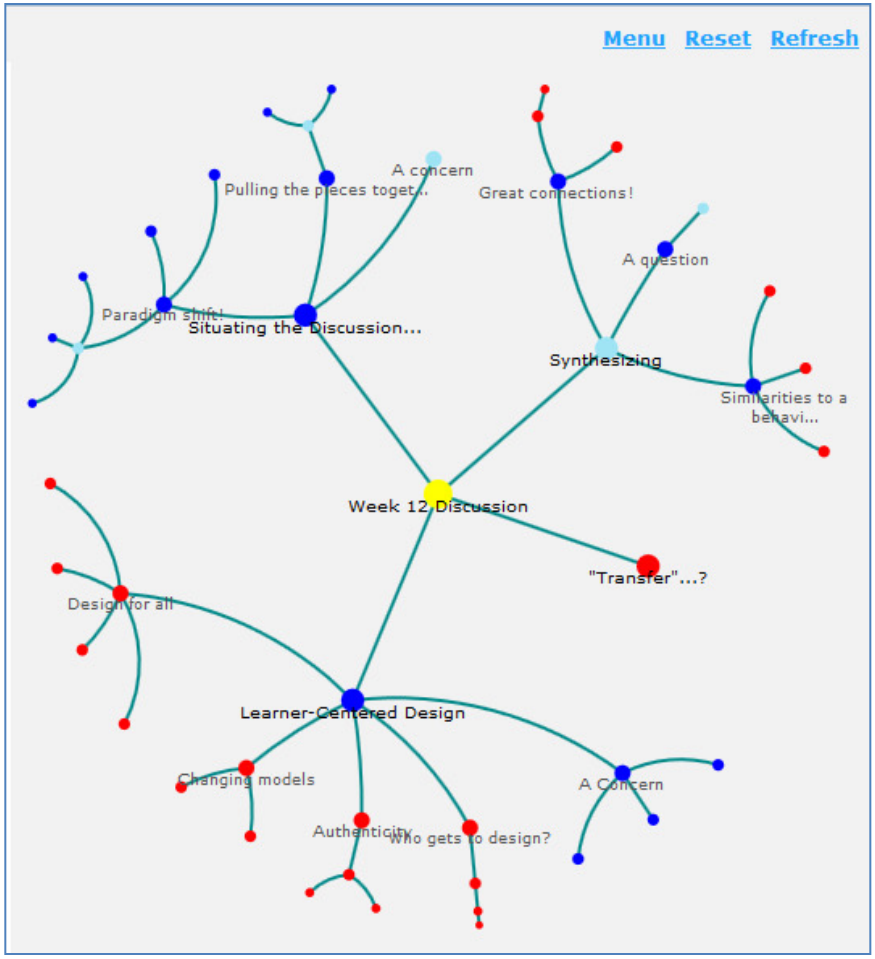
E-Listening Research Project

Discussion Participation Guidelines

Attending to Others Posts

Broad Listening: Try to read as many posts as possible to consider everyone's ideas in the discussion. This can help you examine and support your own ideas more deeply. However, when time is limited it is better to view a portion in depth, then everything superficially.

The visual interface shows posts that you have viewed in **blue and new ones in **red** to help you track this.*



INTEGRATION

Learning Analytics Guidelines

Attending to Others' Posts	
% of posts read	The proportion of posts you read (not scanned) at least once.
It is good to read as many posts as possible to consider everyone's ideas in the discussion However, when time is limited it is better to view a portion in depth, then everything superficially.	

Metric	Me (Week X)	Class (Week X)
Participation range	6 days	5 days
# of sessions	3	11
Average session length	48 min	39 min
% of sessions with posts	67%	49%
# of posts made	4	7
Average post length	386 words	216 words
% of posts read	68%	79%
#of self reviews	2	5
#of peer reviews	12	8

AGENCY
GOAL-SETTING

- Discussion guidelines present metrics as a starting point for consideration, not as absolute arbiters of engagement

- Goal-setting is an explicit and structured part of the learning activity as students set weekly goals for engaging in the online discussions in an online reflection journal (in the LMS)

SAMPLE STUDENT GOALS

"I aim to read all (most) posts [in the discussion], and actively participate in two threads in addition to any I create"

"Well, since I didn't hit last week's goal really I [still] need to do that, also keep the length [of my posts] down and get more interactive with the other kids."

"As a goal for the next discussion, I will try to synthesize ideas from different thread areas"

Agency Reflection

- Establish a rhythm for reflection
  - Weekly cycle of reviewing the analytics
  - Evaluate progress towards the goals
  - Assess when the goals themselves need to be updated or revised

- Provide a dedicated space
  - Online reflective journal (private wiki in the LMS)
  - Supports examination of trajectory over time

Sample Student Reflection

"I found that I wanted the challenge of trying to up the % of overall posts that I reviewed each week. This also meant slowing down my reading since the data would not record a quick read of the information. The overall result was that I think I learned more and was able to get a broader sense of opinion concerning the readings."

Reference Frame

- Continually reminding students of theoretical patterns
- Prompting reflection on individual progress and goals
- Value and danger of comparisons to peers

Sample Mentions of Reference Frames

"I was surprised to see that most of classmates checked the forum more than I did…I also did not expect that they referred [back to] their own post quite many times."

"Since all my numbers are below the average so that makes me feel, 'Oh my gosh, I'm kind of jumping out of this class' or something like that. It is kind of a little bit – sometimes depressing."

"Compared to the previous week, [my] number of reviews of others' posts has been hugely increased … and I did spend more time to read and understand others' posts."

Pilot Study Findings

Integration (technological and pedagogical) made analytics a coherent part of the learning process

Students embraced agency in setting (often recurring) personal goals and evaluating their progress, no "big brother" issues

Individual, peer, and instructor reference frames were important for making sense of the data; reactions were both cognitive and emotional

Reflection on data a powerful starting place
Concrete and proximal goal-setting is harder
Change happens slowly, isn't always intentional, requires support!

Distinct Δ Profiles

Aligned Design

E-Listening Research Project

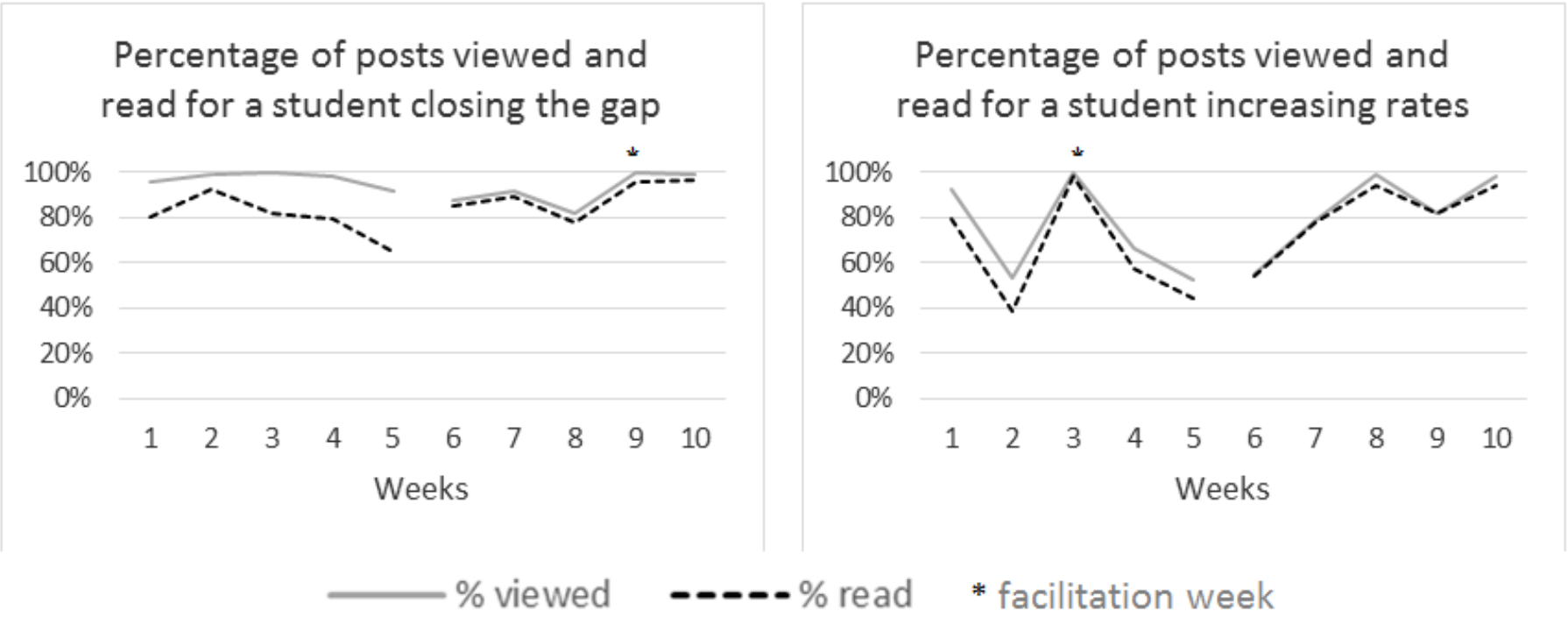


Figure 5. Percentage of posts viewed and read for students who after the introduction of the extracted analytics (a) narrowed the gap between posts viewed and read and (b) raised the percentage of posts both viewed and read

From Data Science to Learning Analytics

Evolution of the E-Listening Research Project

DATA

THEORY

COMPUTATION

DISPLAY

USE

Thank you for listening…

Alyssa Wise

afw3@sfu.ca