The Conceptual and Methodological Future of Large Scale Learning Research

Dr. Alyssa Wise
Associate Professor, Educational Technology & Learning Sciences
Director, Learning Analytics Research Network, New York University
@alywise @NYU_LEARN

NYU | LEARN

MOOCEOLOGY

First, acknowledgement of the important contributions to this presentation and the work on which it is based made by

Yi Cui

Research Assistant

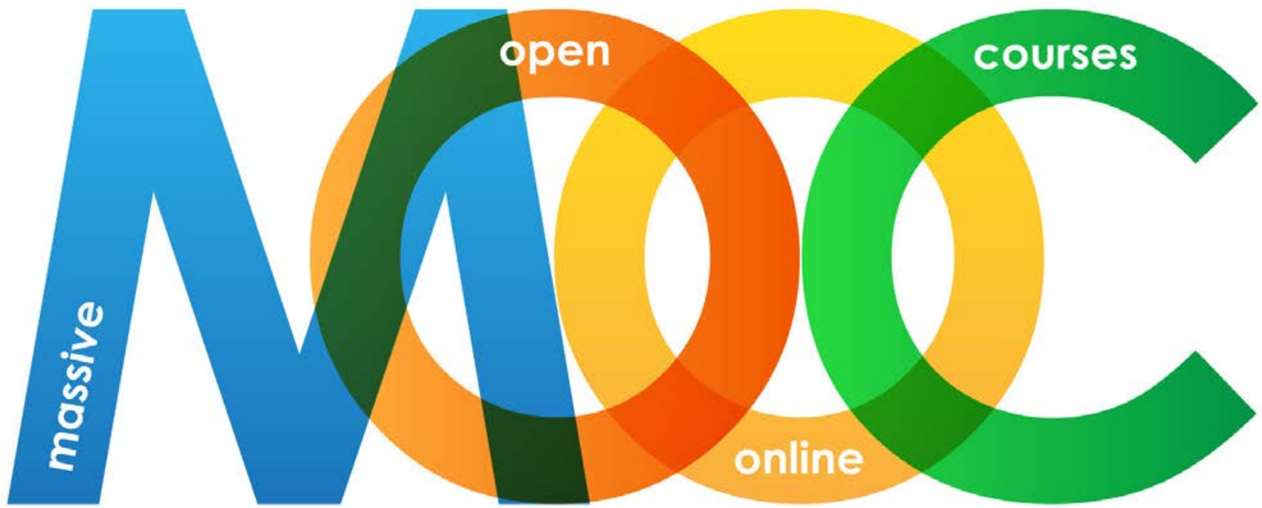
@ NYU-LEARN

Doctoral Candidate

@ Simon Fraser University

Do we have a distinct body of knowledge about MOOC learning?

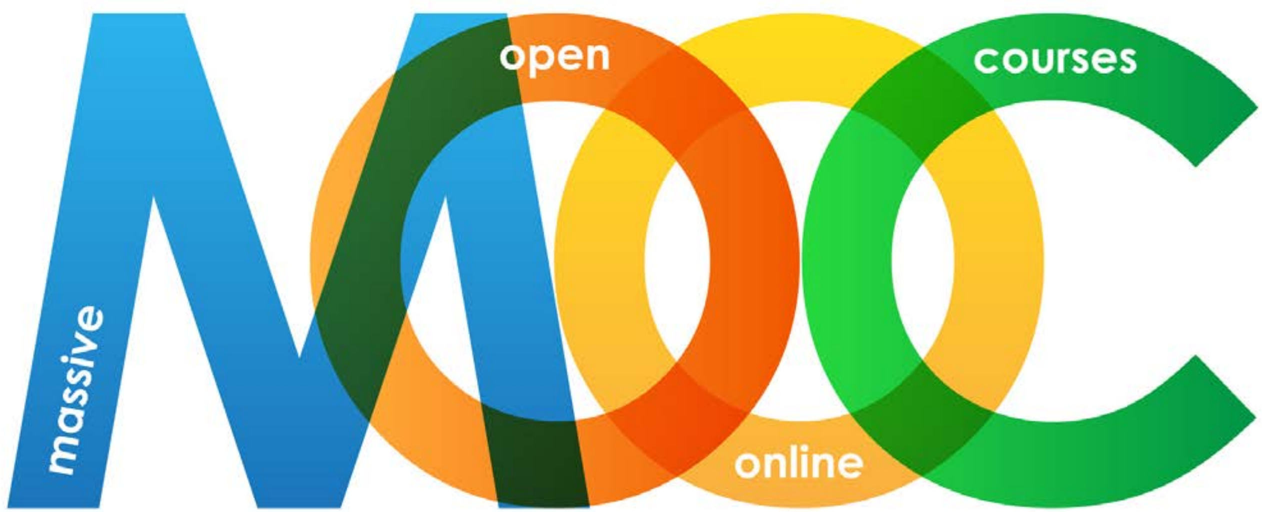
Empirical Findings **Unifying Frameworks**



Overarching Concepts **Unique Qualities**

Why do we need a distinct body of knowledge about MOOC learning?

Empirical Findings **Unifying Frameworks**



Overarching Concepts **Unique Qualities**

MOOCs are different than traditional online courses in important ways

Massiveness

- Large numbers of
  - Learners
  - Activities
  - Courses

- Engender different kinds of pedagogy, opportunities, challenges and interactions
- Require research methods that can be applied efficiently at scale

Openness

- Diversity of
  - Demographics
  - Backgrounds
  - Motivations & Goals

Need to consider

- Subpopulations
- Diverse participation patterns
- New measures of success / definitions of learning

Yet much research remains rooted in traditional paradigms of online learning that do not align with these characteristics

Are studies of retention and grades the most appropriate things to focus on in an open environment which learners come to with diverse backgrounds and goals?

Key Conceptual Questions for MOOC Study

- What are the core characteristics that distinguish MOOCs from other learning environments and thus merit the focus of our attention?
- What different kinds of learning outcomes are valuable and valued in MOOCs?
- What kinds of actions and interactions should be happening in MOOCs (and why)?

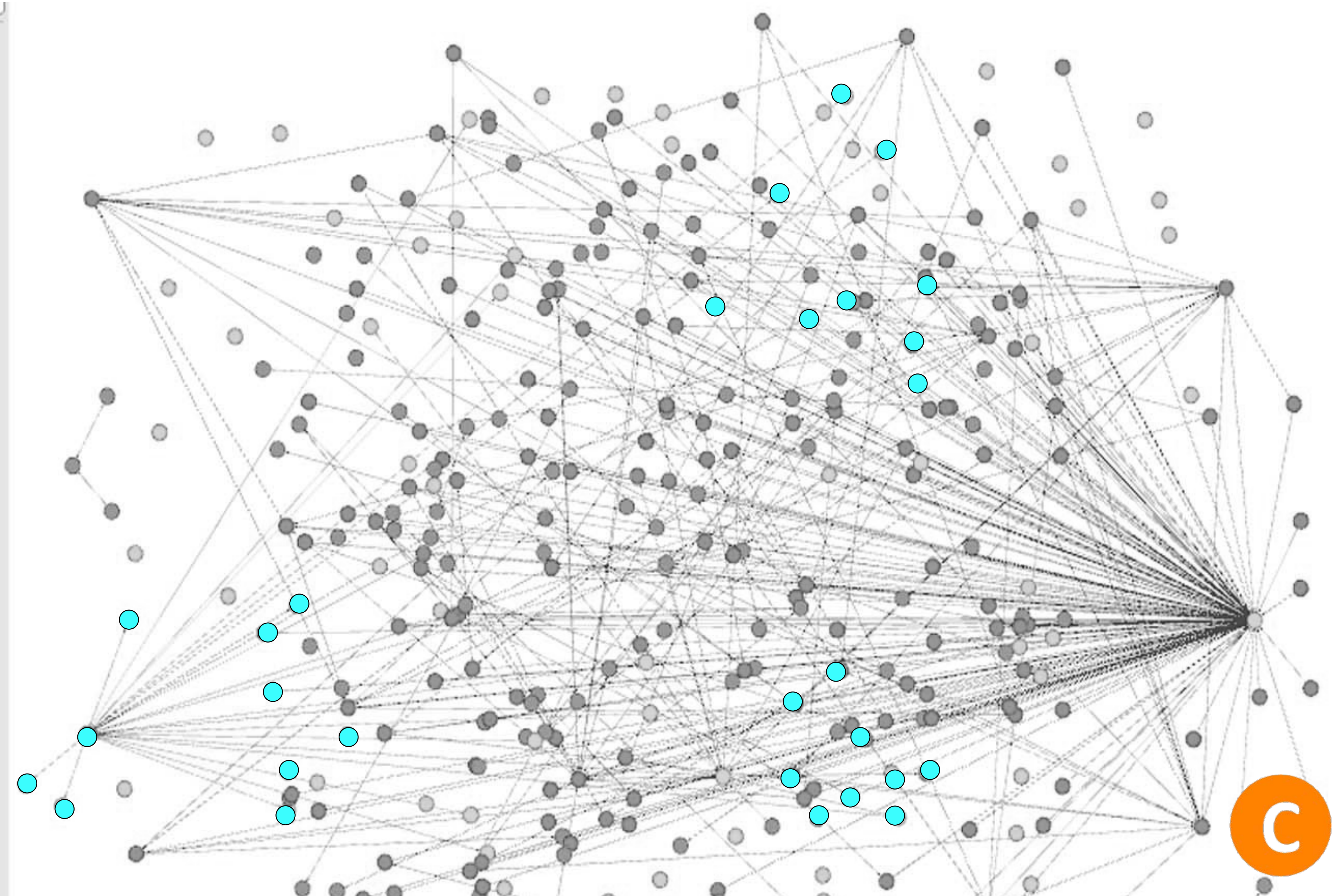
Key Methodological Questions for MOOC Study

- How can and should the power of human intellect and machine computation be brought together to maximize insight?
- How can we handle large quantity of activity efficiently while attending to the complexity of interaction and learning processes?

Investigation of the interaction practices in large-scale learning environments based on analysis of the artifacts left behind by students' and instructors' activity

DATA
THEORY
COMPUTATION
DISPLAY
USE

Conceptualizing the Role of Discussions in Learning: Differentiating Learning-Related & Unrelated Discussions



MOOCeology

Our MOOC Discussion Data

DATA
THEORY
COMPUTATION
DISPLAY
USE

Challenges of post-hoc data

- What is available + sharable
- No data design (settings or structure)
- Limited control (+ info) about learning context of generation
- Ambiguous inferences

Our framing / goals

- "Q&A" style discussions in courses with similar pedagogy
- Increasingly distal generalization
- Consider variations in time

C

MOOCeology

Modelling Questions

DATA

THEORY

COMPUTATION

DISPLAY

USE

1. Do content-related threads in a statistics MOOC discussion forum have distinct linguistic features ?

2. Can these be used to create a model to reliably identify them?

3. Does the model generalize to
   - another offering (same MOOC)?
   - a different statistics MOOC?
   - MOOCs on other topics?

4. Is the model robust over the duration of the course?

MOOCeology

Natural Language Processing

[in Lightside RW]

DATA

THEORY

COMPUTATION

DISPLAY

USE

- Use linguistic features to predict if post is about learning content
- Unit of analysis = Thread
  - Initially represented by Starter post
  - Later, replies incorporated as well
- Hand-coding by (human) research assistants
  - Detailed coding guide + training
  - Good interrater reliability ($\alpha > 0.75$)
- Bag-of-words feature extraction
  - Unigrams and bigrams only, parts of speech unhelpful, stop words *IN*

MOOCeology

Supervised Machine Learning

[in R]

DATA

THEORY

COMPUTATION

DISPLAY

USE

- 2236 extracted features used to train a binary L2 regularized logistic regression model
  - Confusion matrix and data restructuring for model optimization
  - Evaluation via 10-fold cross validation + 4 test sets
- Supplemental Modelling
  - Addition of views + votes
  - Models of only views + votes
  - Tests sets divided into three equal subsets based on time of creation

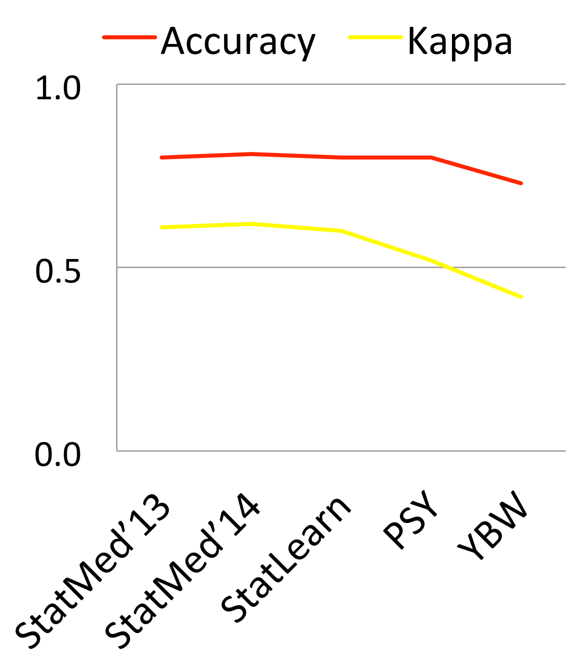
Courses

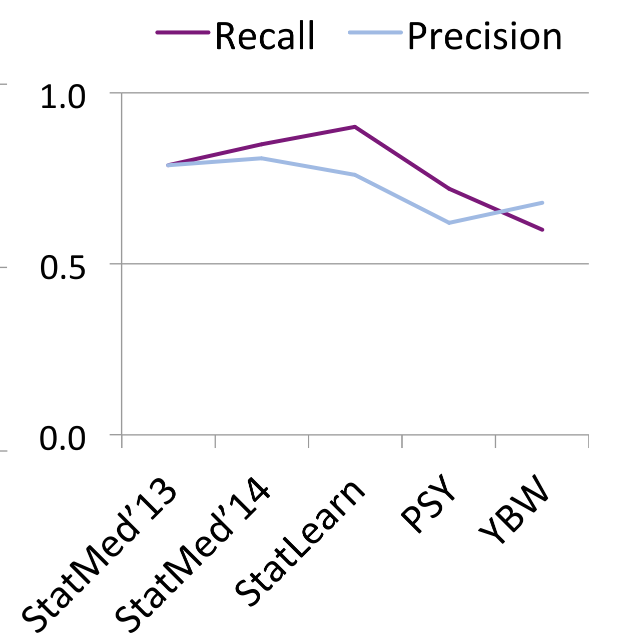
Course	Usage	Total # of Posts	# of Threads (SPs)	# of SPs Coded	SP Content %
StatMed'13 (Statistics)	Training Set	3320	844	837 [844]	47%
StatMed'14 (Statistics)	Test Set: Cross-Offering	1218	310	304 [310]	54%
StatLearn (Statistics)	Test Set: Cross-Course	3030	626	298 [300]	51%
PSY (Psychology)	Test Set: Cross-Domain (Near)	2307	438	438 [438]	28%
YBW (Physiology)	Test Set: Cross-Domain (Far)	2467	825	299[300]	40%

MOOCeology

Model Results

	StatMe d'13	StatMe d'14	Stat Learn	PSY	YBW
Accuracy	0.80	0.81	0.80	0.80	0.73
Kappa	0.61	0.62	0.60	0.52	0.42
Recall	0.79	0.85	0.90	0.72	0.60
Precision	0.79	0.81	0.76	0.62	0.68





DATA

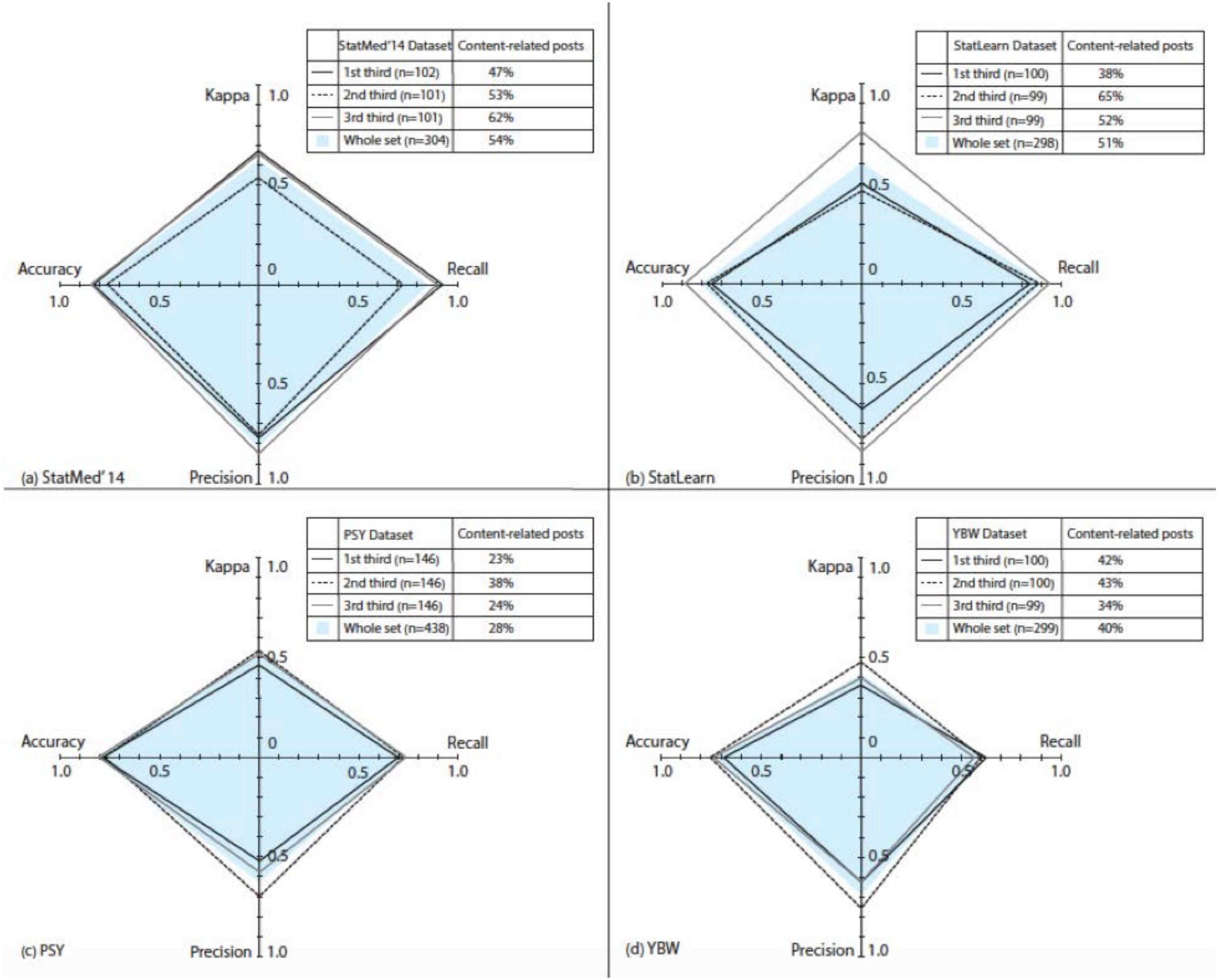
THEORY

COMPUTATION

DISPLAY

USE

Model Performance Across Time Segments



MOOCeology

Improving NLP Classification

D A T A

T H E O R Y

C O M P U T A T I O N

D I S P L A Y

U S E

Dynamic Interrelated Post and Thread Categorization (DIPTiC)

- Verify performance on replies
- Apply classifier to both thread starter and all replies
- Establish cutoff threshold for percent of content replies in content thread
- Compare starter- and reply-based classifications, manual triage on mismatches
- Improvement on StatMed’14 data
  - Accuracy .81 -> .88
  - Kappa .62 -> .76

MOOCeology

DIPTiC in action

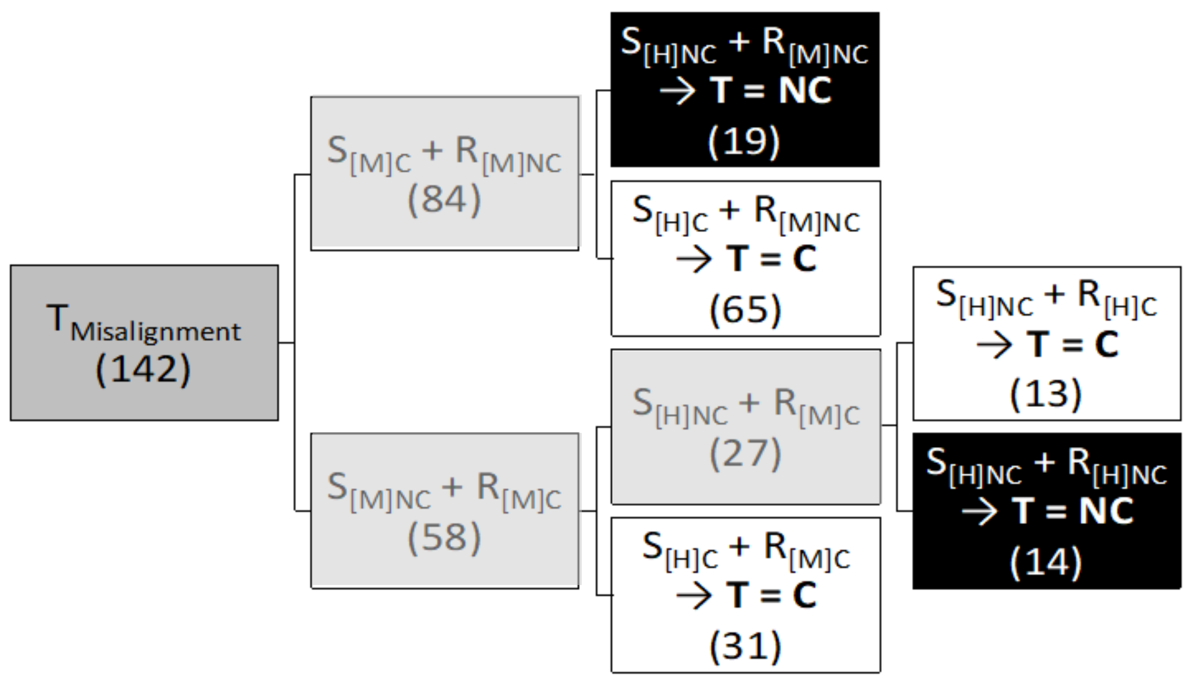
DATA
THEORY
COMPUTATION
DISPLAY
USE

Application to StatMed'14 Data

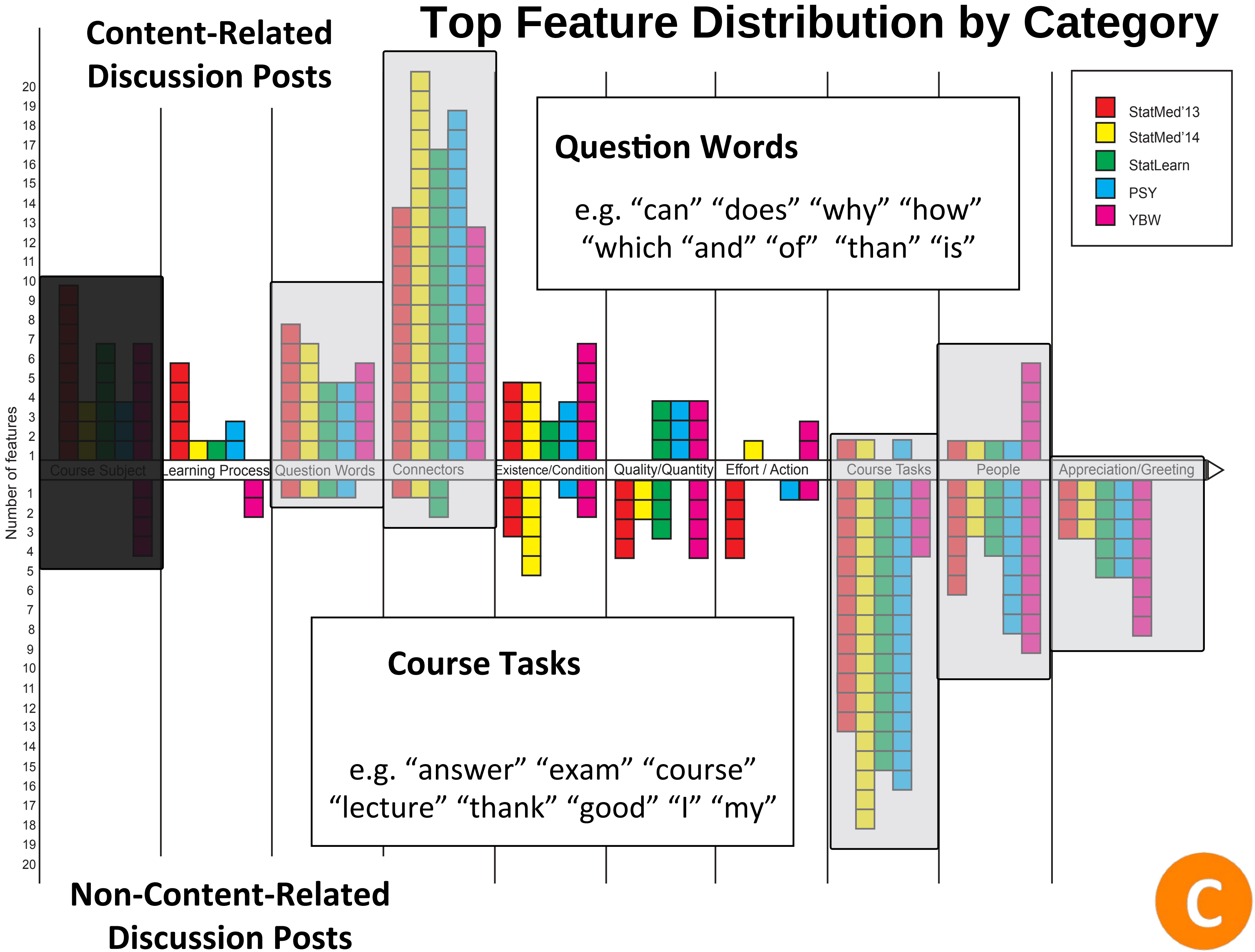
Starter-Reply Mismatches

	$T_{(Reply)}$ = C	$T_{(Reply)}$ = NC
$T_{(Starter)}$ = C	301	**84**
$T_{(Starter)}$ = NC	**58**	257

Recategorization Process



Keys: [M] = Classified by model; [H] = Classified by humans


Top Feature Distribution by Category
Content-Related Discussion Posts
Non-Content-Related Discussion Posts
Number of features
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1
Course Subject
Learning Process
Question Words
Connectors
Existence/Condition
Quality/Quantity
Effort / Action
Course Tasks
People
Appreciation/Greeting
StatMed'13
StatMed'14
StatLearn
PSY
YBW
Question Words
e.g. "can" "does" "why" "how" "which "and" "of" "than" "is"
Course Tasks
e.g. "answer" "exam" "course" "lecture" "thank" "good" "I" "my"

Student & Instructor Support

DATA
THEORY
COMPUTATION
DISPLAY
USE

Post-Hoc Filtering

- Filter to select only content threads
- Reduce # of threads to review by more than half and create > 85% hit rate of those reviewed

Live-Tagging Tool

- Content / non-content label suggested to learners (manual change possible)
- Support student metacognition, awareness of contributions

MOOCeology

Understanding Content Based Interaction

DATA

THEORY

COMPUTATION

DISPLAY

USE

Questions

1. In what ways do unpartitioned, content-related, and non-content social networks show distinct characteristics?

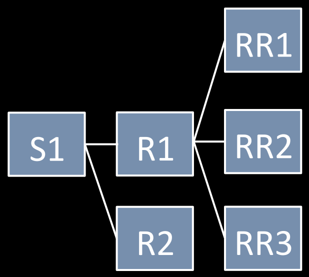
2. What differences in the discussion interactions may account for the distinctions between networks?

3. What effects do different tie definitions have on network characteristics?

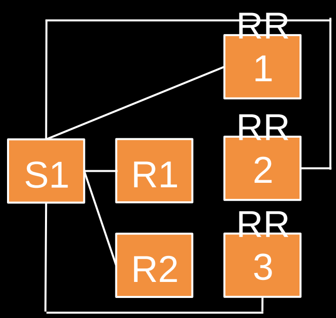
C

MOOCeology

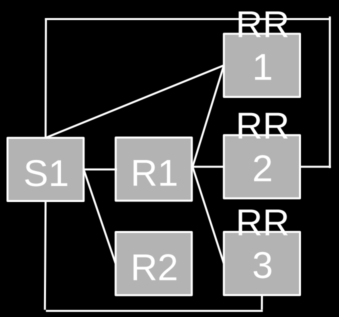
Tie Definition Effects



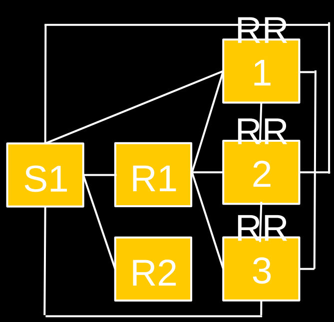
Direct Reply



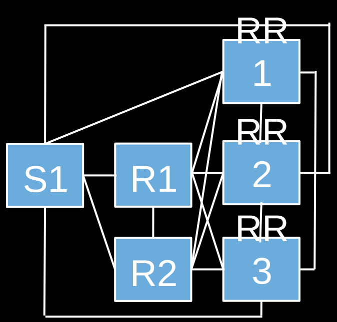
Star



Direct Reply + Star



Limited Copresence



Total Copresence

M

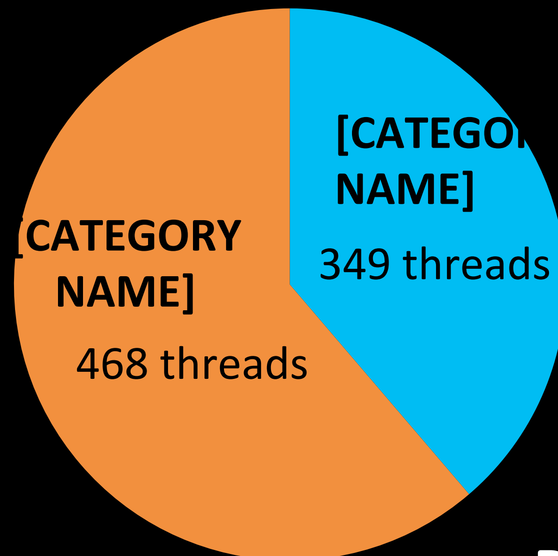
DATA
THEORY
COMPUTATION
DISPLAY
USE

MOOCeology

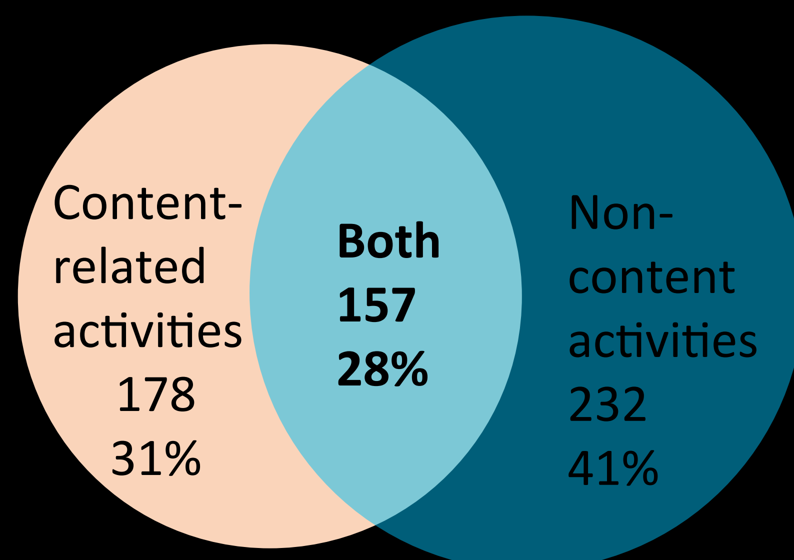
StatMed'14 Profile

DATA
THEORY
COMPUTATION
DISPLAY
USE

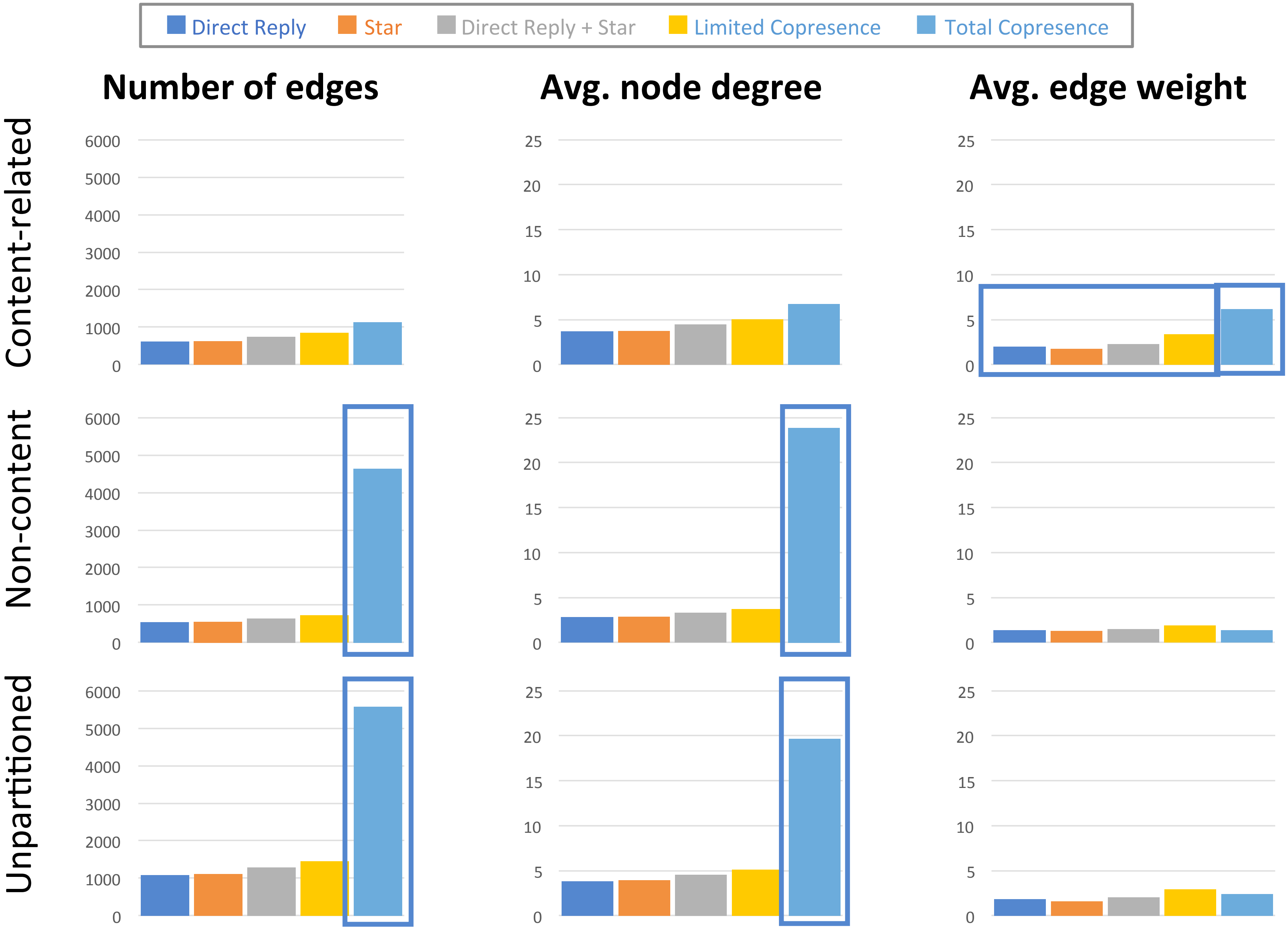
THREADS



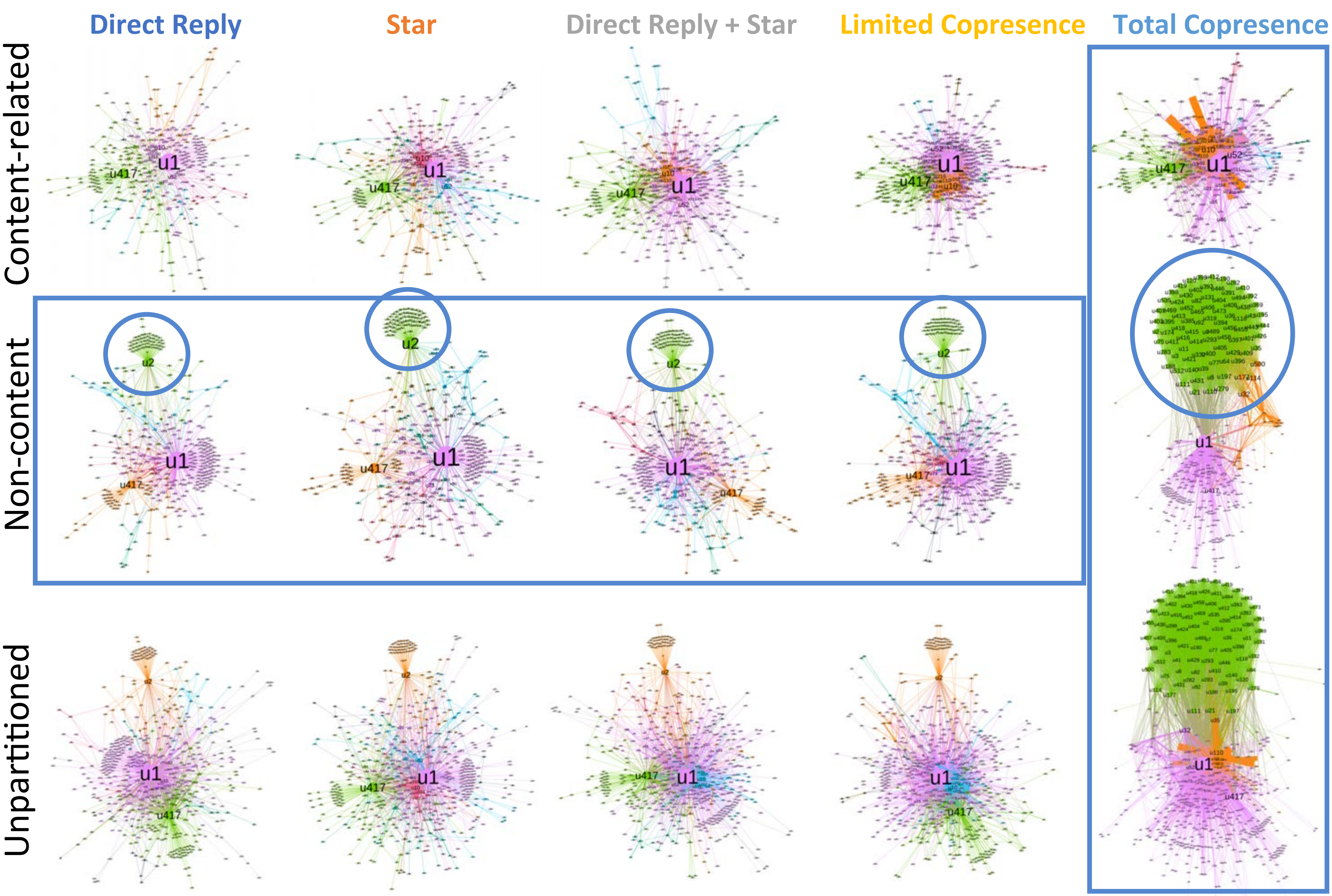
PARTICIPANTS



Tie Definition Effects - Resulting Network Properties



Tie Definition Effects - Resulting Networks



MOOCeology

Qualitative Analysis

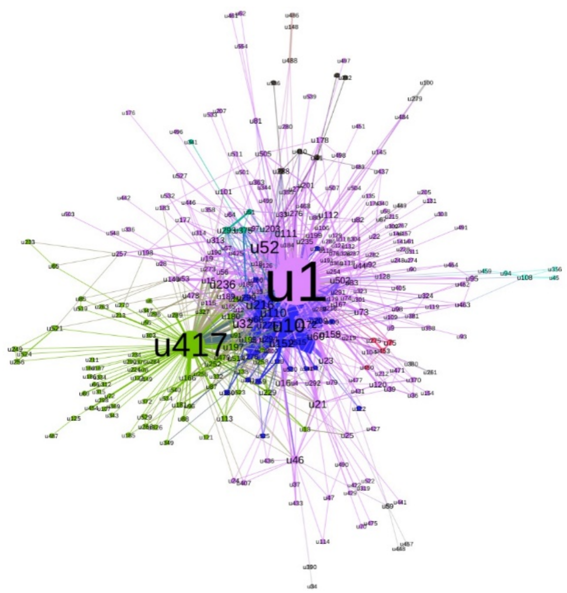
M

DATA
THEORY
COMPUTATION
DISPLAY
USE

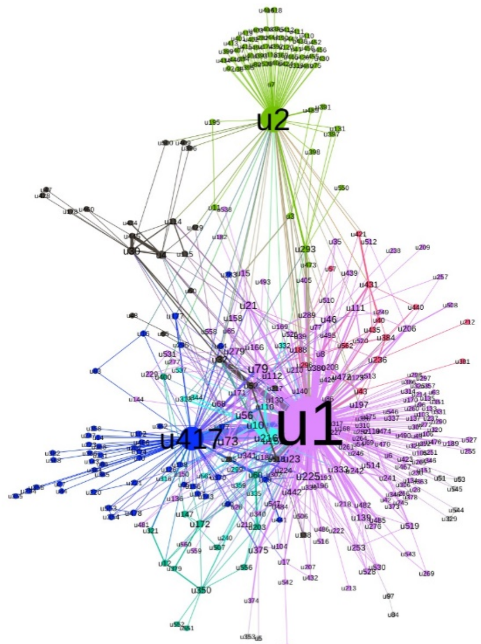
Using Computational Methods to Pinpoint Where to Look

- Identification of relevant communities and sub-networks
- Stratified Random Sampling across thread length to select posts to examine manually
- Again threads taken intact for analysis
- Inductive theme analysis used to make meaning of interactions

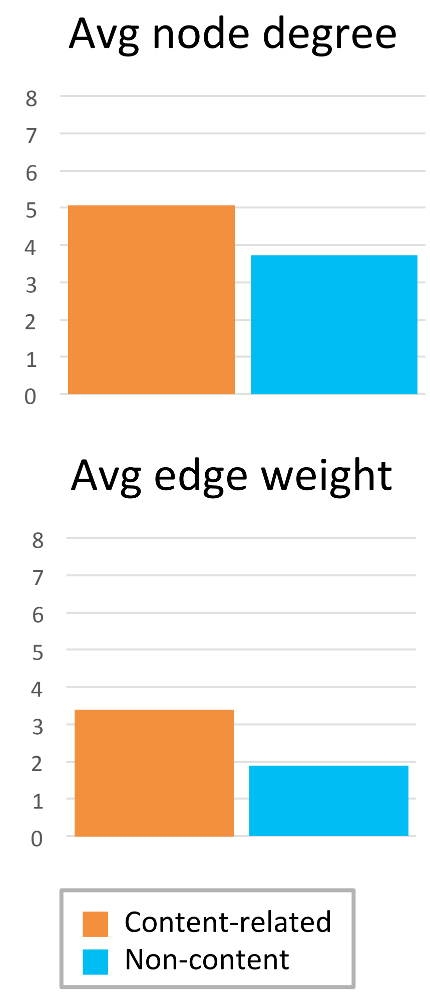
Content vs Non-Content Networks (Limited Copresence Ties)



Content-related network
(# of nodes = 335, # of edges = 848)

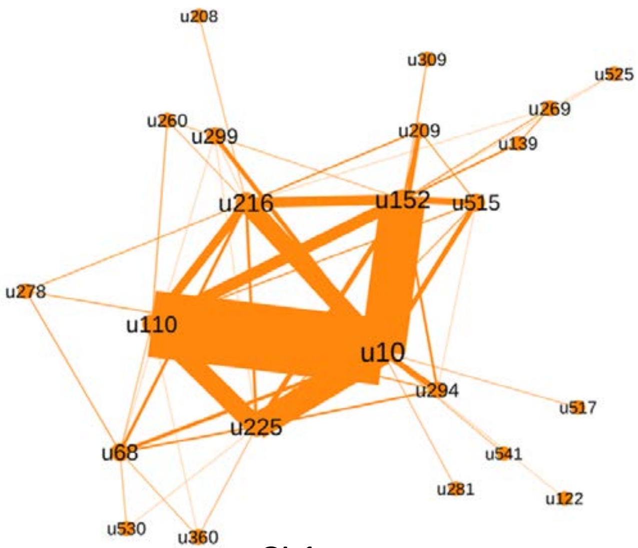


Non-content network
(# of nodes = 389, # of edges = 724)

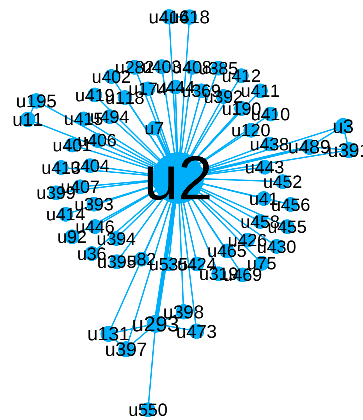


Content interactions had bigger threads with more repeat posters and involved more involved topics, complicated interaction techniques + social presence cues

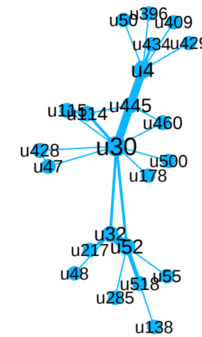
Learner Modules



CL1
(# of nodes = 23, # of edges =57)

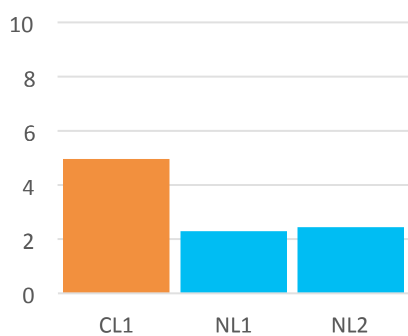


NL1
(# of nodes = 62, # of edges =71)

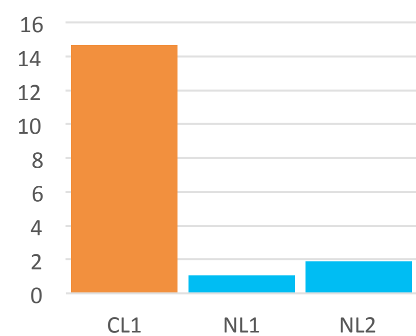


NL1
(# of nodes = 23, # of edges = 28)

Avg node degree



Avg edge weight



CL = Content-related Learner Module
NL = Non-content learner module

Examining Learner Interactions

Across the network content interactions involved more involved topics, complicated interaction techniques + social presence cues

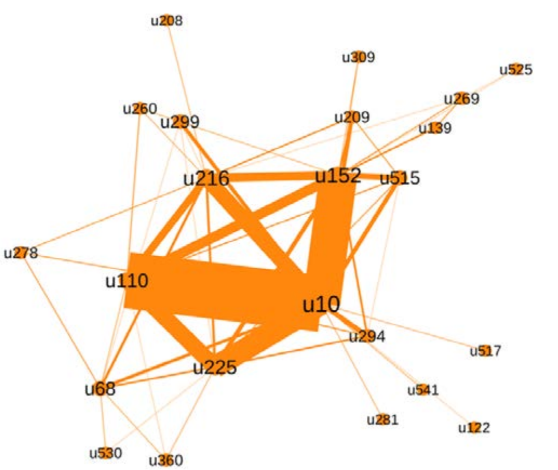
U225*: Congrats [u10]! Yes, it has been hard, but fun, and we learned an awful lot, right?*

U110*: Great! Everyone it was a pleasure to work with you. Thank you….*

U10*: YES [U225]! And [u110] - the test was scary - I thought of my discussion board friends often!!*

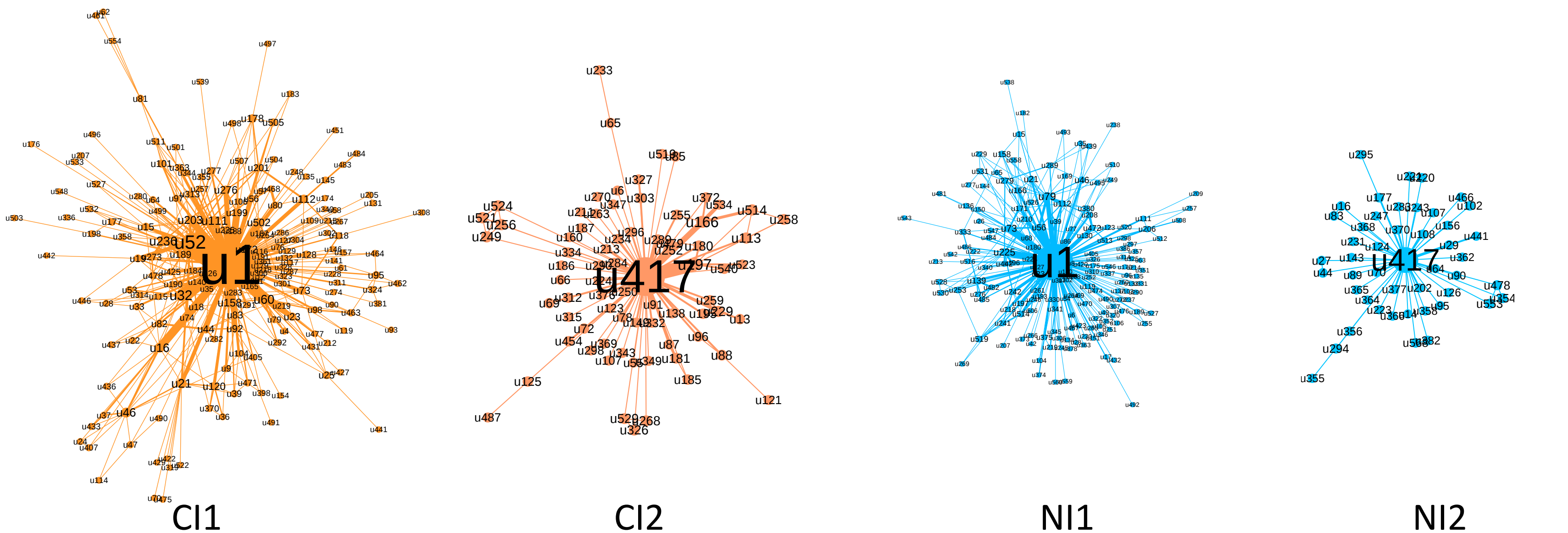
U216*: Thanks, thanks so much to [u10], [u152], [u110], [u225] and everybody who helped us to understand this beautiful course! And in my case also for writing many posts, I see I have improved my English skills and my statistics vocabulary!!!*

U225*: [u10], [u216], [u152], [u110], [u515] and everyone, your discussions helped me so much. I was always a few days behind you in homework - glad I was able to catch up in the last weeks and participate a little bit….*



Content-related learner module 1

Instructor Modules



	CI1	CI2	NI1	NI2
# of nodes (% in network)	184 (54.93%)	75 (22.39%)	168 (43.19%)	47 (12.08%)
# of edges (% in network)	400 (47.17%)	105 (12.38%)	315 (43.51%)	55 (7.6%)
Avg node degree (SD)	4.35 (11.06)	2.8 (7.56)	3.75 (11.18)	2.34 (6.03)
Avg edge weight (SD)	2.23 (3.21)	1.83 (1.72)	2.11 (2.48)	1.20 (0.44)

CI = Content-related instructor module NI = Non-content instructor module

Comparing Instructional Approaches

U1

- Responses at all levels
- Coaching and supporting
- Social presence cues

"Think about it again using the hint and let me know if you have any other questions."

"That is correct - Nice! So how would you use this to solve the question?"

U417

- Responses to thread starters
- Straight forward answers
- Little social presence

"A bell shape is not necessary. You could have a 'bimodal' distribution where the two groups do not follow a bell shape."

MOOCeology

Interactions & Course Performance

DATA
THEORY
COMPUTATION
DISPLAY
USE

Research Gap in the Relationship between Grades / Certificates & Forum Activity

1. Inconsistent findings for which variables are useful predictors for grades / certificates
2. % of variance explained often not reported
3. Little consideration of discussion content

MOOCeology

Interactions & Course Performance

DATA
THEORY
COMPUTATION
DISPLAY
USE

Our Questions

1. Are there differences in MOOC completion and final course grade for learners who did or did not contribute to (content and non-content) discussions?

2. Is forum contribution (measured by quantity and network measures for content and non-content discussions) useful for predicting MOOC course grades?

Learning Context & Data

DATA
THEORY
COMPUTATION
DISPLAY
USE

- Statistics in Medicine MOOC
- 2 Instructors facilitate forums
- Learners
  - 15,073 registered
  - 11,664 with final grade
  - 565 in forum
  - 555 w/ forum + grade data
- 817 threads (inc. 3124 posts) classified as Content / Non-Content using unigram/bigram model + DIPTiC method
- Content, and non-content networks constructed using Limited Copresence tie definition (threshold < 5 replies)

MOOCeology

Interaction & Course Performance

D A T A
T H E O R Y
C O M P U T A T I O N
D I S P L A Y
U S E

Results (Details in LAK'18 Paper)

- Making any kind of forum contribution is associated with a higher likelihood of passing the course, making both kinds is even higher (77% vs 58% vs 32%)
- Contributing to content-related discussions (only) associated with higher final grade, but very small percentage of variance explained
- Network centrality variables don't add anything beyond basic quantity

MOOCeology

Interaction & Course Performance

D A T A

T H E O R Y

C O M P U T A T I O N

D I S P L A Y

U S E

Implications

Three possible explanations for small % of variance in final grade explained by forum

1. Forum participation has little impact on learning.
   → **Need better pedagogical design of discussions**
2. Forum participation is useful, but not measured properly.
   → **Need to assess contribution quality, reading**
3. The type of learning that occurred in the forum is not well captured by final grade.
   → **Need research on alternative perspective on learning, such as over time changes in ways of participation and roles**

Where Do We Look For Learning?

	During Learning	After Learning
Internal Tools	**Assignment & Quizzes** (Brinton et al., 2016; Jiang et al., 2014) **Forum** (Kovanović et al., 2016; Tawfik et al., 2017)	**Final Grades & Certificates** (Bergner et al., 2015; Houston et al., 2017)
External Tools	**Social Media Facebook, Twitter, blogs** (Joksimović et al., 2015a; Joksimović et al., 2015b)	**Github & StackOverflow** (Chen et al., 2016) **Publications & Societies** (Wang, 2017)

Current Work on Alternative Perspectives on MOOC Learning

- Learning is conceptualized as developing one's ability to interact knowledgably in a content domain
- Learning outcomes are then seen as changes in how one positions themselves in a MOOC discussion
- Yi Cui's dissertation work aims to understand position and position taking through combining of content analysis (nature of the contributions made) and social network analysis (nature of one's relation to others)
- She will also examine the impact of pedagogical contexts on interactional processes

How We Addressed Key Conceptual Questions

What core characteristics distinguish MOOCs from other learning environments and merit our attention?

- Mixture of learning-related and unrelated discussions
- Lack of background, context or groupings for interaction

What different kinds of learning outcomes are valuable and valued in MOOCs?

- Traditional learning performance of completion and grades
- Alternative view as ability to interact knowledgably
- Additional perspectives possible

What kinds of actions and interactions should be happening in MOOCs (and why)?

- Questions and connections (focus on material versus self)
- Elaborated threads, repeat engagement, mix social + content

How We Addressed Key Methodological Questions

How can and should the power of human intellect and machine computation be brought together to maximize insight?

- Use computation to identify where in-depth manual analysis is mostly likely to be valuable
- DIPTiC: use multiple measures of computation with humans to resolve discrepancy to make most effective use of people-power
- Use machine learning to extend applicability of human codes

How can we handle large quantity of activity efficiently while attending to the complexity of interaction and learning processes?

- Examine model-identified linguistic features used in context
- Probe intact threads in communities flagged by SNA to generate in-depth understanding of interaction
- Consider conceptual implications of technical decisions (e.g. ties)

Recommendations for the Future of MOOC Research

1. **Conceptualize the diversity** that openness brings as a fundamental aspect of MOOC philosophy, not a problem to be overcome
2. **Examine applicability** of existing online learning theories in the context of massiveness and consider where modification / alternative theories are needed
3. **Combine human** intellect **and machine** computation to probe complex large scale learning processes
4. **Make sense** of **high-level** computational patterns using **low-level** contextualized data; use computational methods to validate small-scale qualitative findings

MOOCEOLOGY

Wise & Cui (in press). **Envisioning a learning analytics for the learning sciences**. *Proceedings of ICLS'18*. ISLS. ** Presenting 11am Wed in Crossover Paper Session at ICLS*

Wise & Cui (online first). **Learning communities in the crowd: Characteristics of content related interactions and social relationships in MOOC forums**. *Computers & Education.*

Wise & Cui (2018). **Unpacking the relationship between discussion forum participation and learning in MOOCs: Content is key**. *Proceedings of LAK'18*. ACM.

Wise, Cui & Jin (2017). **Honing in on social learning networks in MOOC forums: Examining critical network definition decisions**. *Proceedings of LAK'17*). ACM.

Wise et al. (2017) **Mining for gold: Identifying content-related MOOC discussion threads across domains through linguistic modeling**. *Internet and Higher Education, 32,* 11-28.

Cui, Jin & Wise (2017). **Humans and machines together: Improving characterization of large scale online discussions through dynamic interrelated post and thread categorization** (DIPTiC). *Proceedings of Learning at Scale 2017*. ACM.

Wise, Cui & Vytasek (2016). **Bringing order to chaos in MOOC discussion forums with content-related thread identification**. *Proceedings of LAK'16*. ACM.

The Conceptual and Methodological Future of Large Scale Learning Research

Dr. Alyssa Wise
Associate Professor, Educational Technology & Learning Sciences
Director, Learning Analytics Research Network, New York University
@alywise @NYU_LEARN

NYU | LEARN

MOOCEOLOGY