Research directions in learning analytics

Dragan Gašević
@dgasevic

Lorena Barba
@LorenaABarba

Follow

#MOOC research?—terabytes of data on clicks and little understanding of what changed in students' minds, says @bjfr sciencemag.org/content/347/62…

Reich, J. (2015). Rebooting MOOC research - Improve assessment, data sharing, and experimental design. *Science, 347(6217),* 30-31, http://bit.ly/1s3b5kS

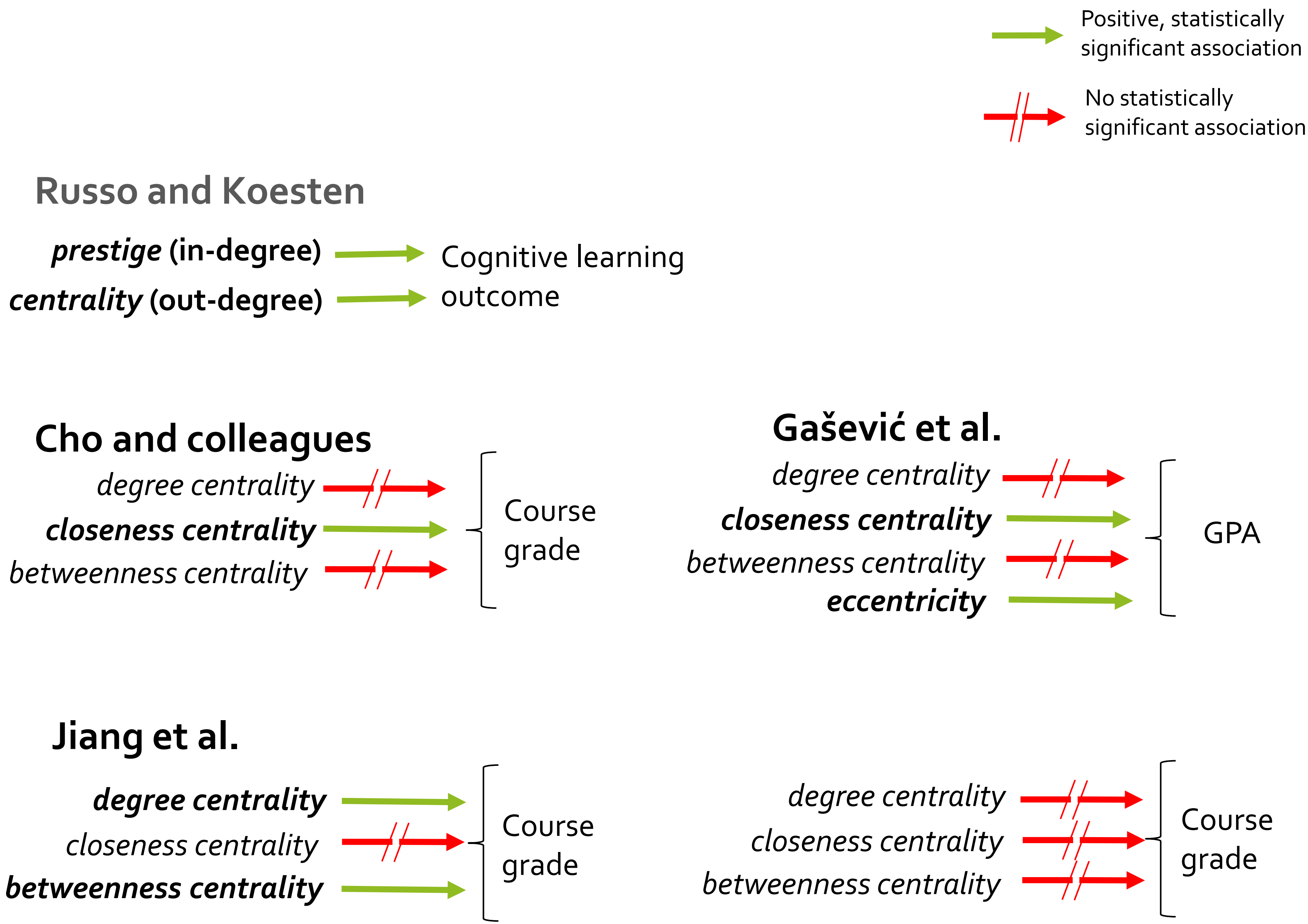
THEORY INFORMED LEARNING ANALYTICS

Counts don’t count much if decontextualized

Wilson, T.D. (1999). Models in information behaviour research. *Journal of Documentation*, 55(3), 249 – 270.

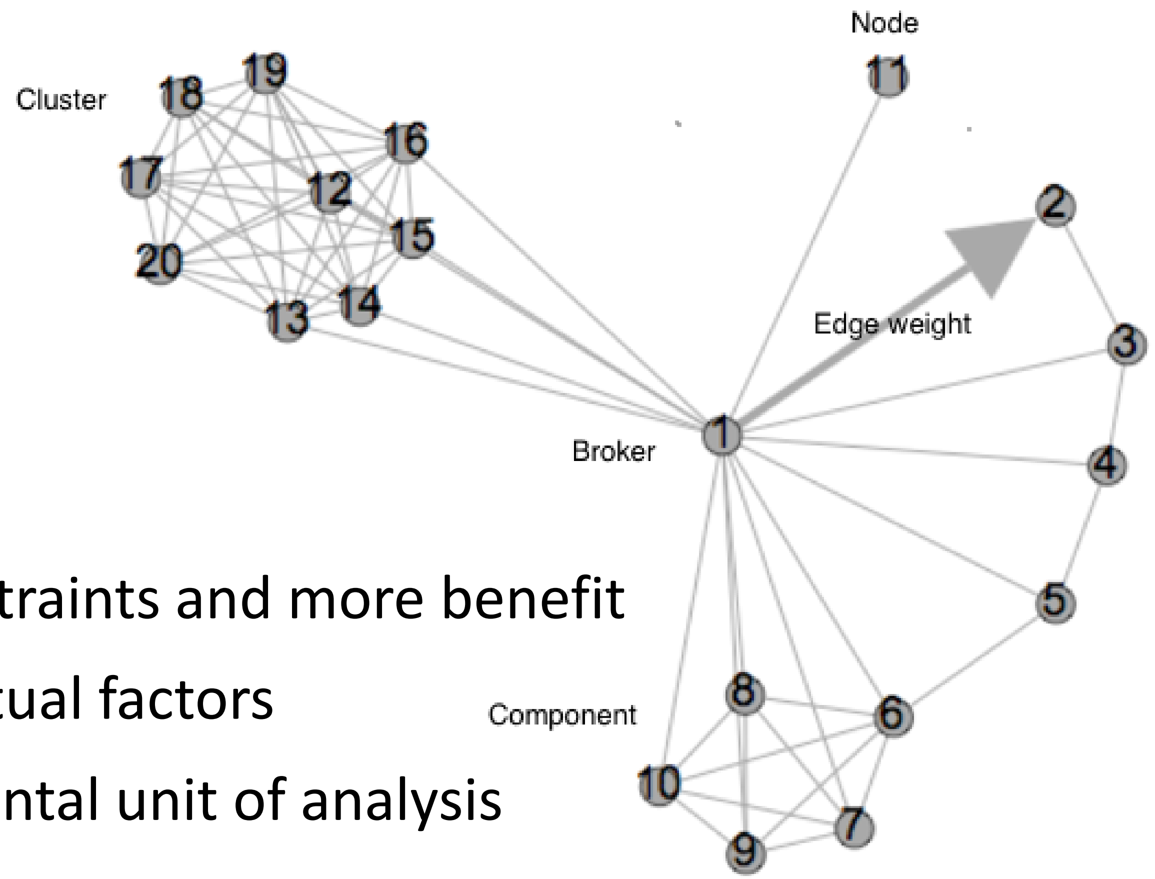
How do strong and weak effect translating network position into performance?

Joksimović, S., Manataki, A., Gašević, D., Dawson, S., Kovanović, K., de Kereki, I. F., "Translating network position into performance: Importance of Centrality in Different Network Configurations," *In Proceedings of the 5th International Conference on Learning Analytics & Knowledge (LAK 2016)*, Edinburgh, Scotland, UK, 2016 (in press).



Joksimović, S., Manataki, A., Gašević, D., Dawson, S., Kovanović, K., de Kereki, I. F., "Translating network position into performance: Importance of Centrality in Different Network Configurations," *In Proceedings of the 5th International Conference on Learning Analytics & Knowledge (LAK 2016)*, Edinburgh, Scotland, UK, 2016 (in press).

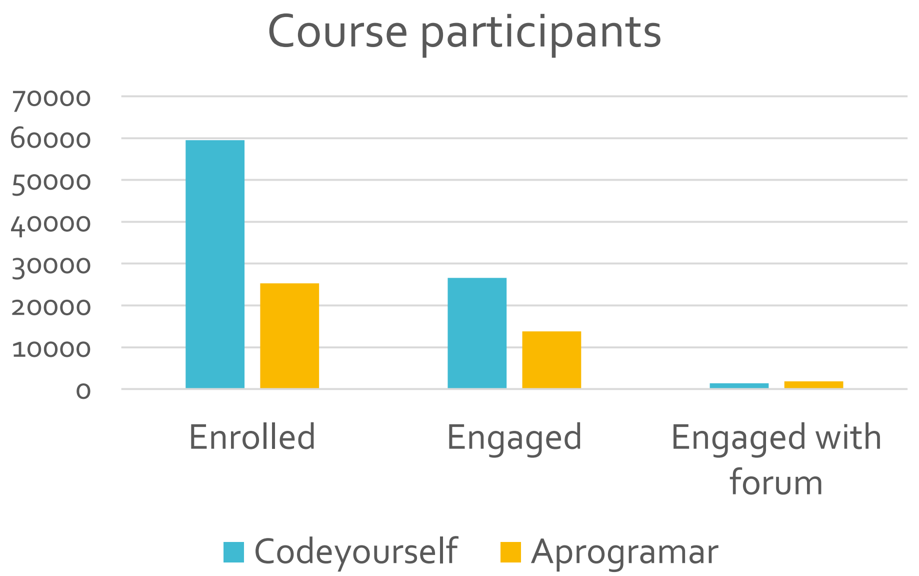
Simmel's theory of social interactions

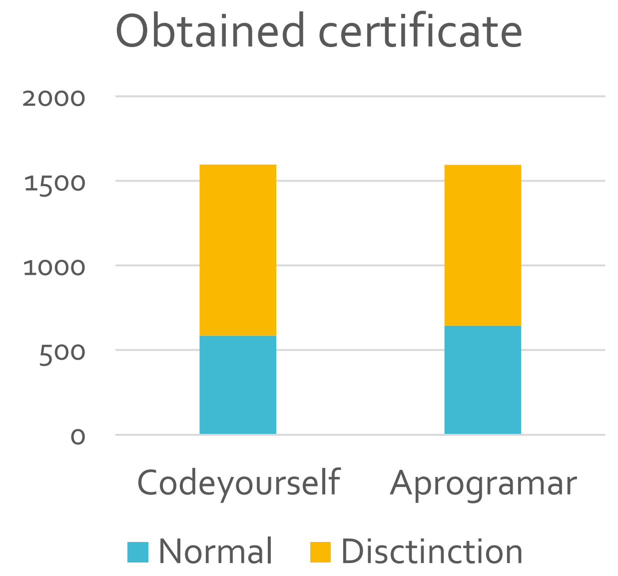


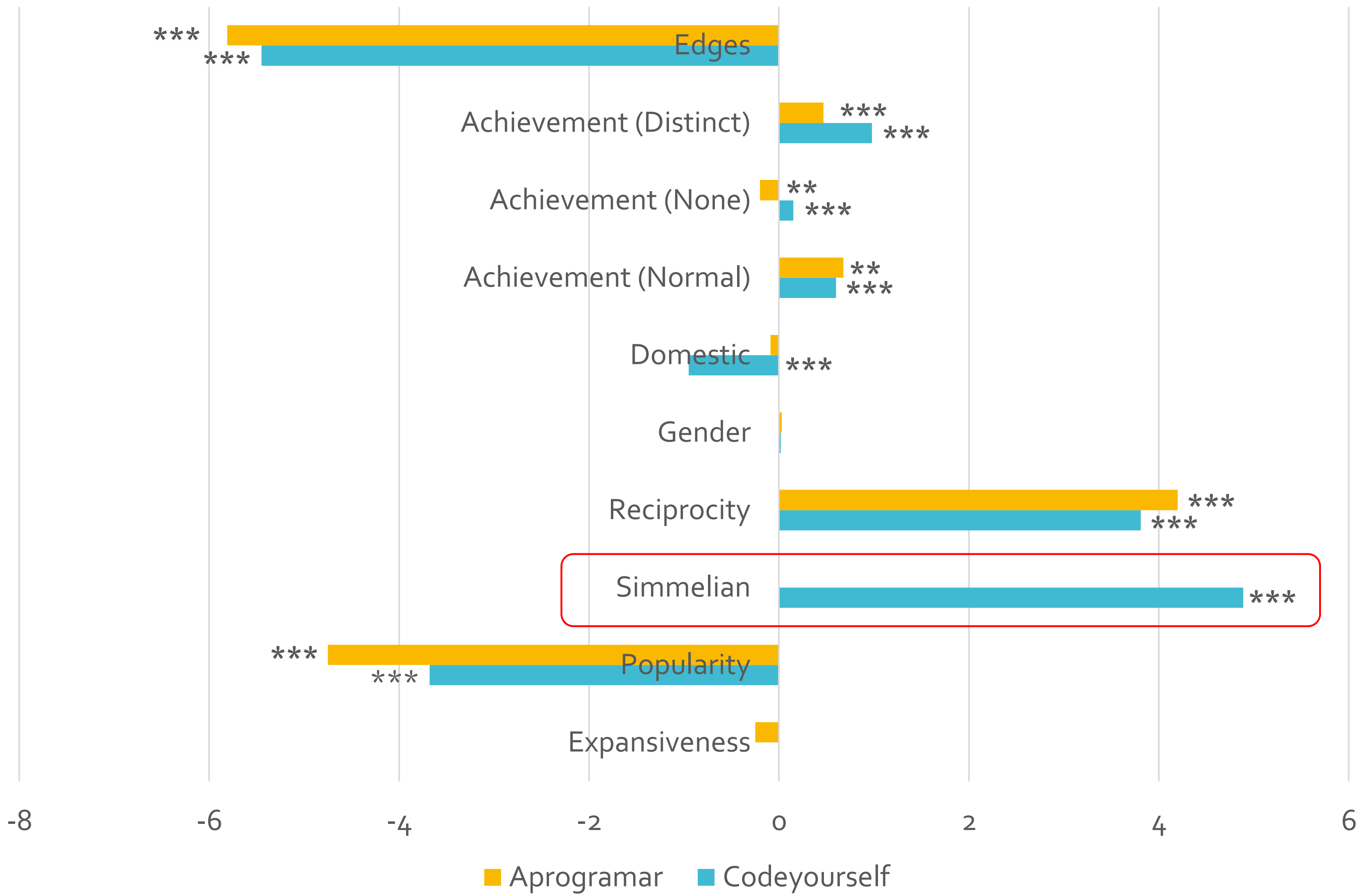
- Centrality $\not\Rightarrow$ less constraints and more benefit
- Importance of contextual factors
- Triads as the fundamental unit of analysis

Study

- Platform: Coursera
- Courses: Code Yourself! (English), ¡A Programar! (Spanish)
- Certificate: 50% for the coursework; 75% - distinction

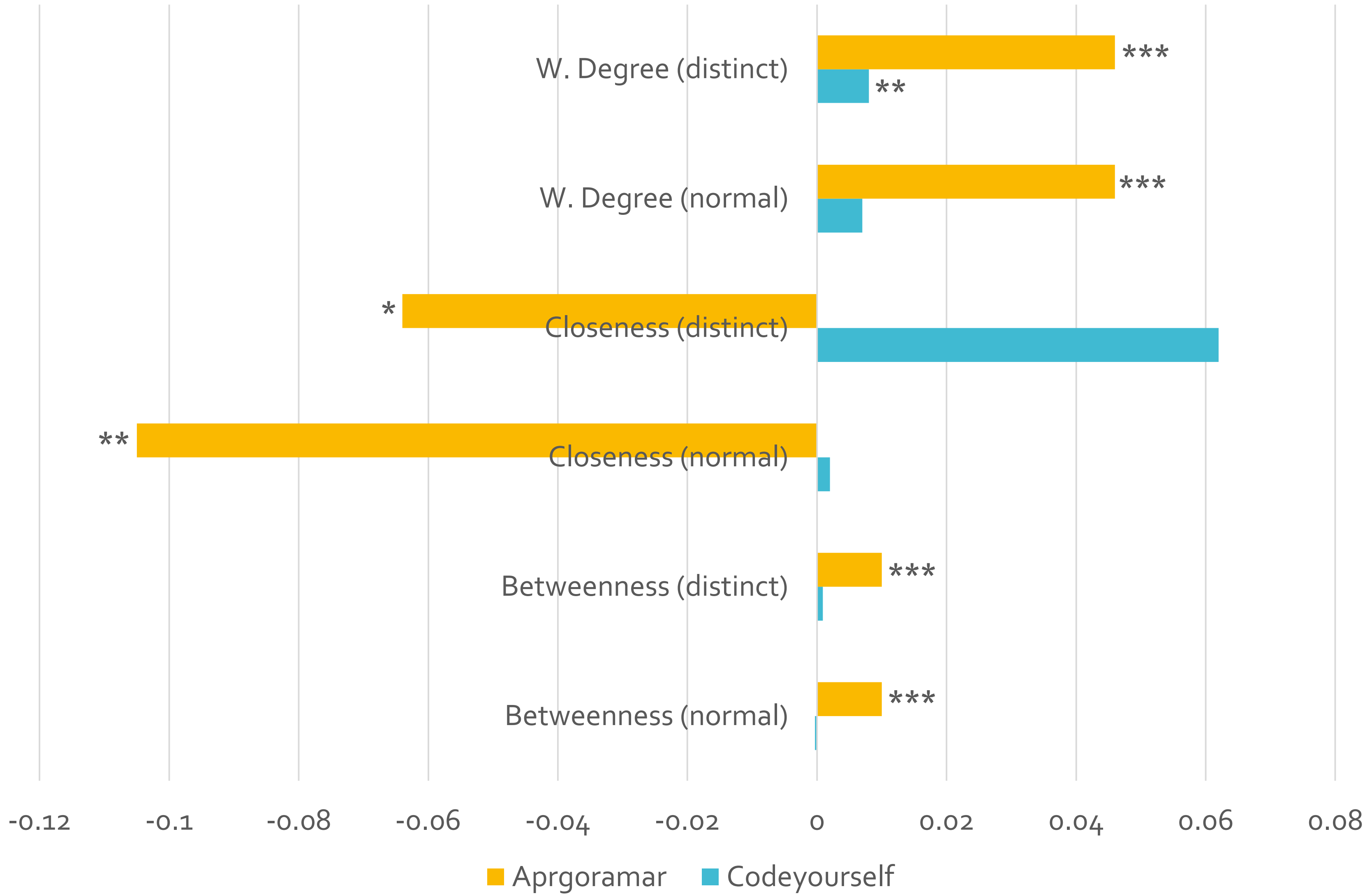






Analysis of the estimates for the two ERG models

Note: * p<.05; ** p<.01; *** p<.001



Multinomial regression analysis – network centrality(independent) and course completion (dependent

Note: * p<.05; ** p<.01; *** p<.001

In order to provide meaningful visualizations, estimates for betweenness centrality were multiplied by 100 (only for the presentation purposes)

Learning analytics is about learning

Learners construct knowledge

Learners are agents

Winne, P. H. (2006). How software technologies can improve research on learning and bolster school reform. *Educational Psychologist*, *41*(1), 5–17.

Learning analytics is about learning

Conditions, Operations, Products, Evaluation, Standards (COPES)

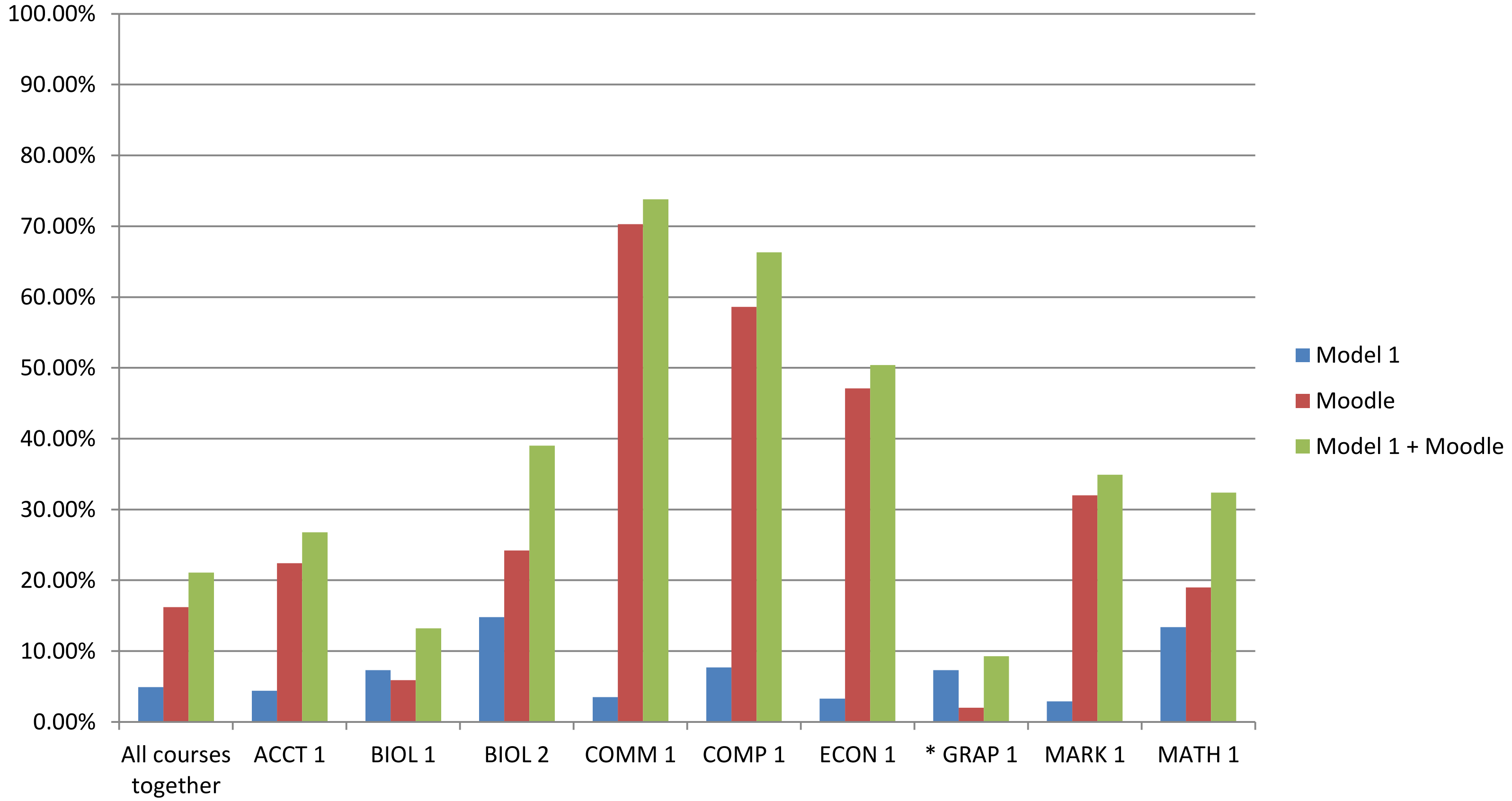
Winne, P. H. (1997). Experimenting to bootstrap self-regulated learning. *Journal of Educational Psychology, 89*(3), 397-410.

Learning context

Instructional conditions shape learning analytics results

Gašević, D., Dawson, S., Rogers, T., Gašević, D. (2016). Learning analytics should not promote one size fits all: The effects of course-specific technology use in predicting academic success. *The Internet and Higher Education, 28*, 68–84.

Predictive Power Diversity



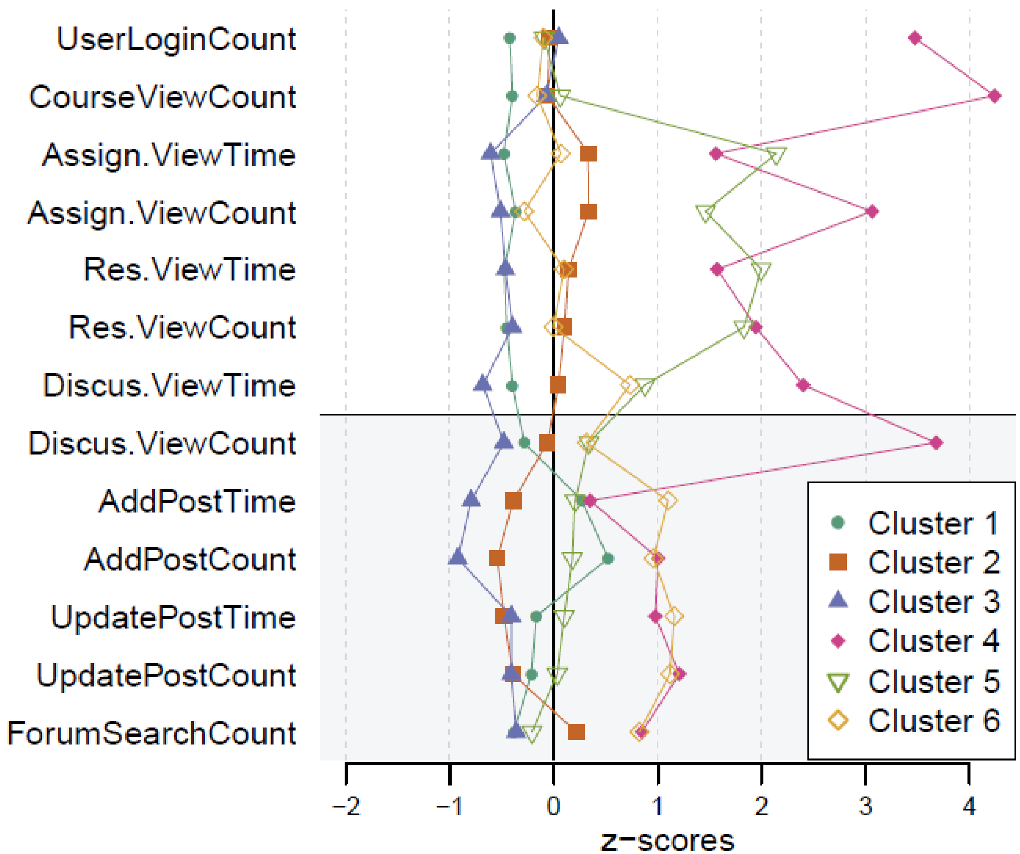
Model 1 – demographic and socio-economic variables
* - not statistically significant

Learner agency

More time online does not always mean better learning

Kovanović, V., Gašević, D., Joksimović, S., Hatala, M., Adesope, S. (2015). Analytics of Communities of Inquiry: Effects of Learning Technology Use on Cognitive Presence in Asynchronous Online Discussions. *The Internet and Higher Education*, 27, 74–89.

Learner profiles – use of LMS



Large effect sizes (1.4-2.5 σ) on critical thinking and academic success

Kovanović, V., Gašević, D., Joksimović, S., Hatala, M., Adesope, S., “Analytics of Communities of Inquiry: Effects of Learning Technology Use on Cognitive Presence in Asynchronous Online Discussions,” *The Internet and Higher Education*, 2015 (submitted).

PROCESS NATURE OF LEARNING

How students study with technology?

Categorization

Deep and surface approaches to learning

Trigwell, K., & Prosser, M. (1991). Relating approaches to study and quality of learning outcomes at the course level. *British Journal of Educational Psychology*, *61*(3), 265-275.

Significant role of instructions on approaches to learning

Trigwell, K., Prosser, M., & Waterhouse, F. (1999). Relations between teachers' approaches to teaching and students' approaches to learning. *Higher Education, 37*(1), 57–70.

Effects of students' own decisions

Internal conditions
(e.g., metacognition and motivation)

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-Regulated Learning: Beliefs, Techniques, and Illusions. *Annual Review of Psychology, 64*, 417-444. doi:10.1146/annurev-psych-113011-143823

Student profiling

Unsupervised approaches

Lust, G., Elen, J., & Clarebout, G. (2013). Students' tool-use within a web enhanced course: Explanatory mechanisms of students' tool-use pattern. *Computers in Human Behavior*, *29*(5).

Sequences of activities

Sequence or process mining, HMMs, etc.

Reimann, P., Markauskaite, L., Bannert, M. (2014). e-Research and learning theory: What do sequence and process mining methods contribute? *British Journal of Educational Technology, 45*(3), 528-540.

What learning strategies do students follow while using technology?

Do learning strategies of students change over time while using technology?

Context

Year one engineering course in computer systems at University of Sydney

Enrolment: 300 students

One lecture (2 hours) + one tutorial (2 hours) + one lab (3 hours)

Assessment: midterm + final + project

Flipped classroom with 100% digital content

Next topic

Course Organization

Enter terms for a quick search

Go

ELEC1601 Computer Systems Weekly Schedule

Course Organization		Course Objectives	
Week 1: Organization and Computer System Overview	Week 2: Information Encoding	Week 3: Computer Memory	Week 4: Boolean algebra and combinatorial logic
Week 5: Sequential circuit design	Week 6: Midterm Exam	Project Description	
Week 7: AVR Architecture	Week 8: AVR Instruction Set Architecture	Week 9: Assembly Programs	Week 10: Addressing Modes
Week 11: High Level Programming Constructs	Week 12: Subroutines	Week 13: Exam simulation	

Table Of Contents

Previous topic

2.2.14. Overflow and Underflow in Floating Point Encoding

Next topic

2.2.18. Encoding multiple numbers

Enter terms for a quick search

Go

Week 3

To do

- Prepare tutorial answering the problem sequence about data encoding.
- Give a quick read to the material to cover in the lab. Comment it with your team mates.
- Watch the videos, read the material, and complete the problem sequence to prepare the lecture.

Assessment: Lecture preparation (video and sequence of problems) 1 mark, Tutorial preparation and participation (sequence of problems and participation in session) 1 mark.

To know

- How to analyze the capacity of two encoding schemes and detect advantages and disadvantages.
- Encode arbitrary set of data in a digital system.
- Control the light sensor and the buzzer with the Arduino UNO board.
- How memory is organized in a computer system and how is data stored.

Tutorial: Exercises on Information Encoding

Lab: Lights and Buzzer

Lecture: Computer Memory

Table Of Contents

Previous topic

2.2.14. Overflow and Underflow in Floating Point Encoding

Next topic

2.2.18. Encoding multiple numbers

Enter terms for a quick search

Go

- Watch the videos, read the material, and complete the problem sequence to prepare the lecture.

Assessment: Lecture preparation (video and sequence of problems) 1 mark, Tutorial preparation and participation (sequence of problems and participation in session) 1 mark.

To know

- How to analyze the capacity of two encoding schemes and detect advantages and disadvantages.
- Encode arbitrary set of data in a digital system.
- Control the light sensor and the buzzer with the Arduino UNO board.
- How memory is organized in a computer system and how is data stored.

Tutorial: Exercises on Information Encoding

Lab: Lights and Buzzer

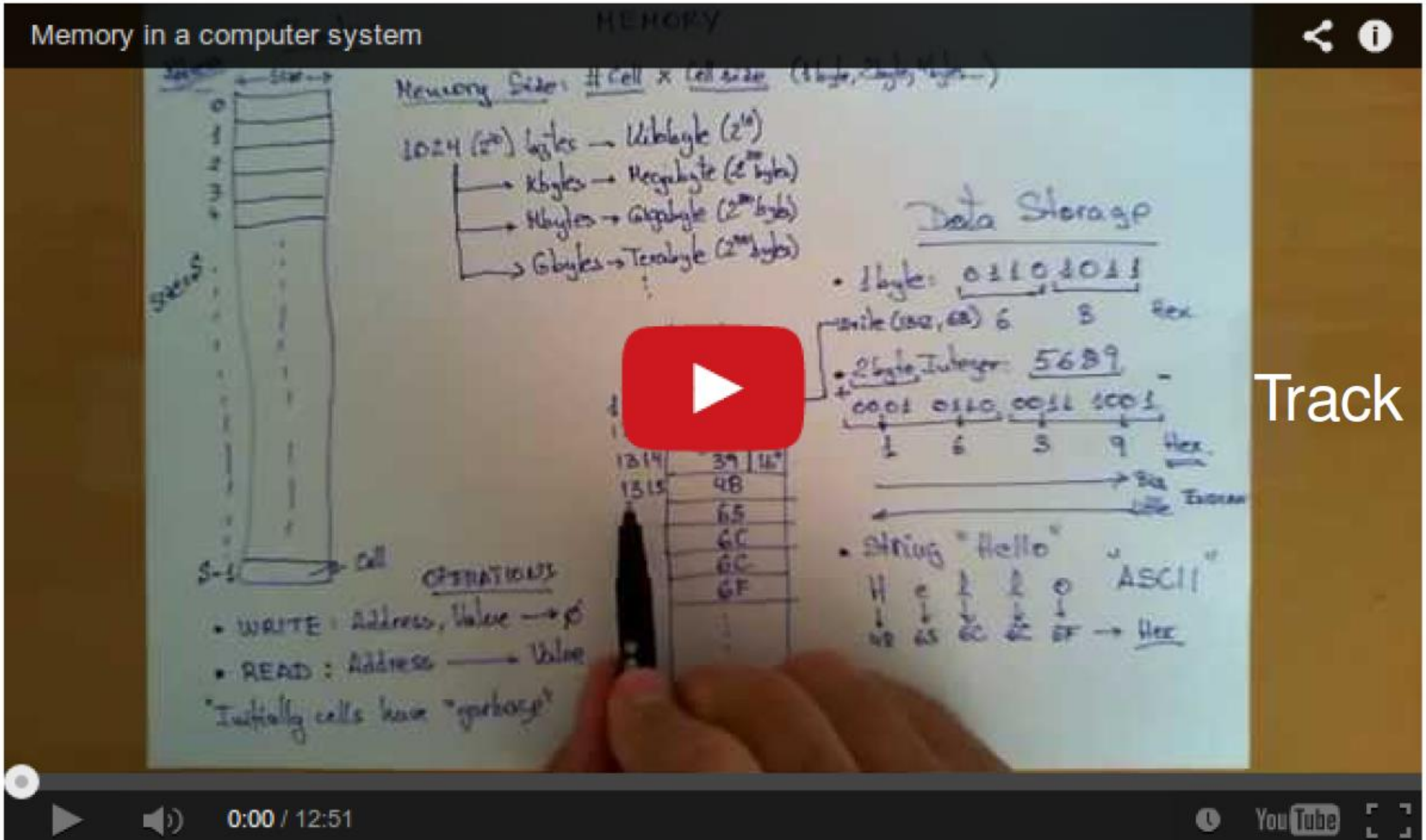
Lecture: Computer Memory

Track

- Activities to do **before the session**:
  - VIDEO: Encoding Sets of Symbols
  - 3.2.1. VIDEO: The structure and operations in memory
  - 3.2.3. Read about how data types are stored in memory
  - 3.2.5. VIDEO: How tables/arrays are stored in memory
  - 3.2.8. VIDEO: Memory Indirection Video
  - Sequence of problems about memory storage (**score to be added to the course marks**).
  - Print and bring to the lecture the **`Week 3 Lecture Worksheet`**.
- Activities to do **during the session** (you do not have to work on them before that time):
  - 2.2.17. Encoding Colors
  - 3.2.2. Before and after memory operations
  - 3.2.4. Store two integers in memory
  - 3.2.6. Access to array elements
  - 3.2.9. Indirection to three integers
- Do you need to review the entire material for this session? Go to *How to store data in memory*

3.2.1.2. Resources

- The video summarizing how memory works and data is stored.



- **The annotations produced during the video**

Track

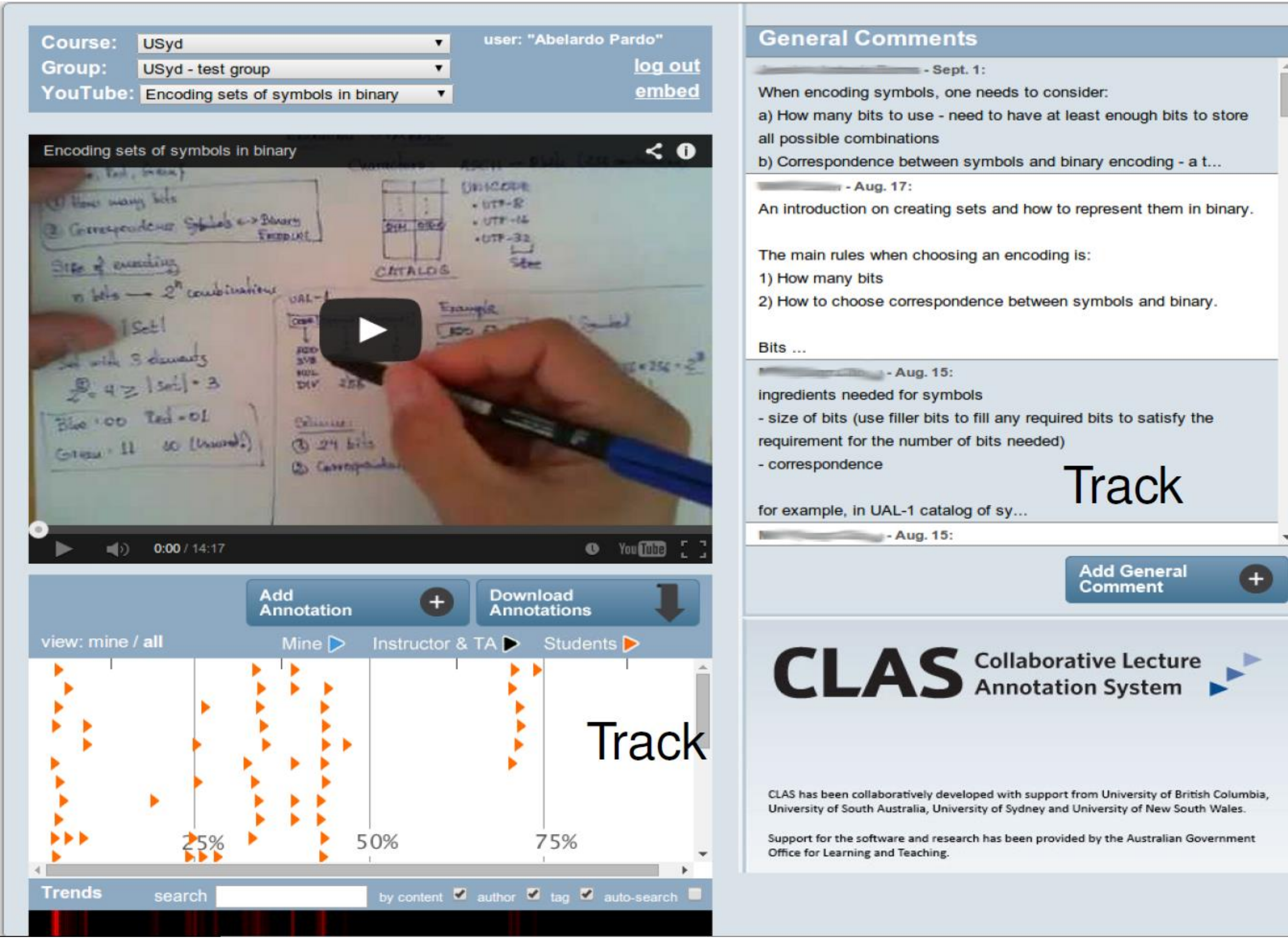

Course: USyd
Group: USyd - test group
YouTube: Encoding sets of symbols in binary
user: "Abelardo Pardo"
log out
embed
Encoding sets of symbols in binary
0:00 / 14:17
Add Annotation
Download Annotations
view: mine / all
Mine
Instructor & TA
Students
Track
25%
50%
75%
Trends
search
by content
author
tag
auto-search
General Comments
- Sept. 1:
When encoding symbols, one needs to consider:
a) How many bits to use - need to have at least enough bits to store all possible combinations
b) Correspondence between symbols and binary encoding - a t...
- Aug. 17:
An introduction on creating sets and how to represent them in binary.
The main rules when choosing an encoding is:
1) How many bits
2) How to choose correspondence between symbols and binary.
Bits ...
- Aug. 15:
ingredients needed for symbols
- size of bits (use filler bits to fill any required bits to satisfy the requirement for the number of bits needed)
- correspondence
for example, in UAL-1 catalog of sy...
Track
- Aug. 15:
Add General Comment
CLAS
Collaborative Lecture Annotation System
CLAS has been collaboratively developed with support from University of British Columbia, University of South Australia, University of Sydney and University of New South Wales.
Support for the software and research has been provided by the Australian Government Office for Learning and Teaching.

Table Of Contents

Previous topic

3. The Memory in a Computer System

Next topic

3.2. Activities about Data Representatino in Memory

Enter terms for a quick search

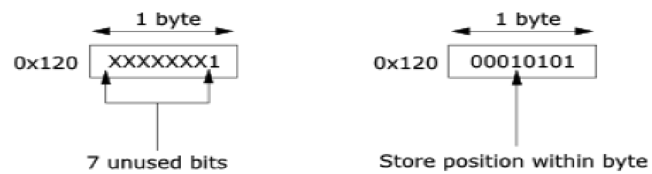
Go

short	Integer	16 bits	[-32768, 32767]
int	Integer	32 bits	[-2147483648, 2147483647]
long	Integer	64 bits	[-9223372036854775808, 9223372036854775807]
float	IEEE-754 Floating point	32 bits	[±1.4012985E-45, ±3.4028235E+38]
double	IEEE-754 Floating point	64 bits	[±4.94065645841246544E-324, ±1.797693134862315 7E+308]

The simple rule to store data in memory is to use as many consecutive cells as needed to store a complex data structure. The address of the first cell from which the data structure is stored will be referred to as *the data address*. Analogously, when a data structure is stored at a certain memory address, what it really means is that it is stored with as many memory cells as required starting at the given address.

3.1.4.1. Storage for booleans

Booleans, despite of being the most simple data structure (only two possible values) are not the easiest ones to store. Memories allow to access the values stored in a cell (typically 1 byte minimum), so storing a single bit means accessing information that is not directly available and requires additional processing to extract or insert the appropriate value. With this technique, eight boolean values can be stored in a 1 byte memory cell. However, the disadvantage of this technique is that to access the boolean, we need to know the memory address and *the position of the bit inside the byte*, which is a number between 0 and 7. As an alternative, booleans can be stored in an entire memory cell leaving the rest of bits untouched. In this case, all bits but one are wasted, but the access to the value is much faster. The following figure shows these two possibilities in a memory with 1 byte cells.



Two strategies to store booleans

Question 1 If a program has to store 128 booleans, how many bytes are required if eight booleans are stored per byte?

- A. 128
- ✖ B. 64
- C. 32
- D. 16

Track

Again Solutions

Question 2 How many **bits** are wasted (that is, not used to store anything) if the 128 booleans of the previous question are stored in one boolean per byte?

Table Of Contents

Previous topic

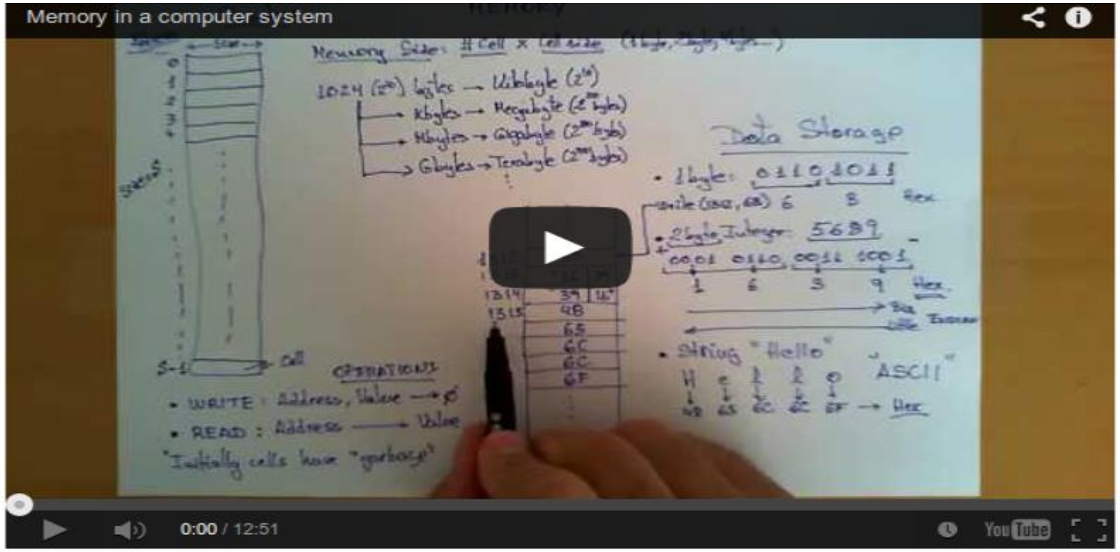
VIDEO: Encoding Sets of Symbols

Next topic

3.2.3. Read about how data types are stored in memory

Enter terms for a quick search

Go



- **`The annotations produced during the video`**

3.2.1.3. Workplan

1. Watch the video
2. Answer the following questions

Question 1 If a memory has cells of 8 bits each, and a total of 1024 cells, the size of that memory is

A. 1024 bits
B. 8 Kilobytes
C. 1024 Kilobytes
✓ D. 1 Kilobyte

Again | Solutions

Track

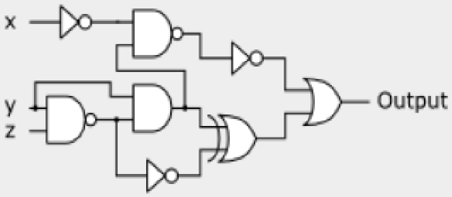
Question 2 The number that uniquely identifies a cell of memory is

✗ A. its content
B. its address
C. its value
D. none of them

Again | Solutions

Circuit Evaluation

Consider the following combinational circuit



Track

Can you determine the value of the output of the XOR gate without knowing the input value of x?

A. Yes, it is zero
B. Yes, it is one
C. Yes, if I know the values of y and z.
D. No, I need to know all three input signals

Your answer? ○ A ○ B ○ C ○ D

Grade

Your score: 17.65%

Next exercise

Features

ACE	number of times a student expanded a part of the course page with exercise or a problem to solve
EQT.CO	the number of correctly solved multiple choice questions (MCQs) embedded in the lecture materials
EQT.IN	number of incorrectly solved MCQs embedded in the lecture materials
EQT.SH	number of times the student requested to see solution to MCQs embedded in the lecture materials
EXC.CO	number of correctly solved exercises/problems
EXC.IN	number of incorrectly solved exercises/problems
VEQ.CO	number of correctly solved MCQs associated with the course videos
VEQ.IN	number of incorrectly solved MCQs associated with the course videos
VEQ.SH	number of times the student requested to see solution to MCQs associated with the course videos
VID.PA	number of times the course videos were played
VID.PL	number of times the course videos were paused

Analysis

Agglomerative hierarchical clustering
based on weekly data

Latent class analysis for course pathways

Hidden Markov models (multinomial)

Results – Week 2

N = 272

attributes	Median; (Q1, Q3)	Median; (Q1, Q3)	Median; (Q1, Q3)	Median; (Q1, Q3)
cluster	1	2	3	4
freq	64	41	69	98
ACE	8; (3, 15.25)	44; (32, 68)	22; (16, 36)	33; (22, 54)
EQT.CO	0; (0, 0)	6; (1, 32)	0; (0, 3)	7; (2.25, 12.75)
EQT.IN	0; (0, 0)	4; (0, 15)	0; (0, 1)	4.5; (1, 8.75)
EQT.SH	0; (0, 0)	1; (0, 7)	0; (0, 0)	2; (0, 4)
EXC.CO	0; (0, 0)	38; (38, 38)	19; (19, 19)	19; (19, 19)
EXC.IN	0; (0, 0)	25; (14, 36)	17; (11, 28)	15; (10, 21)
VEQ.CO	0; (0, 4.5)	12; (6, 17)	0; (0, 3)	10; (6, 12)
VEQ.IN	0; (0, 1.25)	5; (3, 10)	0; (0, 1)	5; (3, 7)
VEQ.SH	0; (0, 0)	4; (1, 8)	0; (0, 1)	3; (1, 5)
VID.PA	0; (0, 0)	5; (1, 16)	0; (0, 4)	7; (2, 16.75)
VID.PL	0; (0, 0)	8; (1, 17)	1; (0, 5)	8; (3, 20)
SC_MT_TOT	12; (10, 15)	15; (11, 18)	14; (11, 16)	15; (12, 17)
SC_FE_TOT	15; (10.75, 20)	16; (11, 27)	17; (14, 25)	19; (14, 29)

Cluster 1: Disengaged (64, 23.5%)
Cluster 2: The Most Engaged (41, 15.1%)
Cluster 3: Exercise-focused (69, 25.4%)
Cluster 4: Engaged and high-performing (98, 36%)

Results – Week 6, pt. 1 of 2

N = 287

attributes	Median; (Q1, Q3)	Median; (Q1, Q3)	Median; (Q1, Q3)	Median; (Q1, Q3)	Median; (Q1, Q3)
cluster	1	2	3	4	5
freq	55	33	49	113	37
ACE	90; (62, 111.5)	68; (50, 96)	49; (36, 71)	30; (13, 41)	50; (32, 63)
EQT.CO	66; (51, 94.5)	38; (16, 57)	10; (1, 23)	2; (0, 12)	13; (3, 21)
EQT.IN	33; (24, 45.5)	22; (11, 47)	5; (1, 18)	2; (0, 6)	7; (3, 11)
EQT.SH	6; (2.5, 16.5)	5; (1, 27)	2; (0, 11)	0; (0, 3)	2; (0, 6)
EXC.CO	66; (29.5, 114.5)	77; (58, 100)	6; (4, 26)	4; (4, 15)	76; (49, 101)
EXC.IN	40; (10.5, 70)	60; (44, 69)	8; (3, 30)	5; (1, 10)	49; (30, 67)
VEQ.CO	5; (0, 16)	44; (30, 60)	15; (9, 27)	0; (0, 4)	0; (0, 4)
VEQ.IN	4; (0, 8)	24; (18, 40)	12; (7, 17)	0; (0, 3)	0; (0, 2)
VEQ.SH	0; (0, 2)	7; (1, 16)	6; (2, 12)	0; (0, 1)	0; (0, 1)
VID.PA	7; (1, 18.5)	31; (14, 61)	28; (10, 68)	1; (0, 4)	2; (0, 10)
VID.PL	8; (1, 17.5)	31; (13, 83)	29; (13, 74)	1; (0, 6)	3; (0, 15)
SC_MT_TOT	15; (13, 17)	15; (11, 16)	13; (10, 15)	13; (10, 16)	16; (13, 17)
SC_FE_TOT	20; (15, 30.5)	17; (14, 26)	14; (11, 19)	17; (12, 24)	18; (14, 29)

Cluster 1: Highly engaged, exhibiting 'guessing' behavior, focused on MCQs in lecture materials (55, 19.1%)
Cluster 2: Highly engaged, exhibiting 'guessing' behavior, focused on video-related activities (33, 11.5%)

Results – Week 6, pt. 1 of 2

N = 287

attributes	Median; (Q1, Q3)	Median; (Q1, Q3)	Median; (Q1, Q3)	Median; (Q1, Q3)	Median; (Q1, Q3)
cluster	1	2	3	4	5
freq	55	33	49	113	37
ACE	90; (62, 111.5)	68; (50, 96)	49; (36, 71)	30; (13, 41)	50; (32, 63)
EQT.CO	66; (51, 94.5)	38; (16, 57)	10; (1, 23)	2; (0, 12)	13; (3, 21)
EQT.IN	33; (24, 45.5)	22; (11, 47)	5; (1, 18)	2; (0, 6)	7; (3, 11)
EQT.SH	6; (2.5, 16.5)	5; (1, 27)	2; (0, 11)	0; (0, 3)	2; (0, 6)
EXC.CO	66; (29.5, 114.5)	77; (58, 100)	6; (4, 26)	4; (4, 15)	76; (49, 101)
EXC.IN	40; (10.5, 70)	60; (44, 69)	8; (3, 30)	5; (1, 10)	49; (30, 67)
VEQ.CO	5; (0, 16)	44; (30, 60)	15; (9, 27)	0; (0, 4)	0; (0, 4)
VEQ.IN	4; (0, 8)	24; (18, 40)	12; (7, 17)	0; (0, 3)	0; (0, 2)
VEQ.SH	0; (0, 2)	7; (1, 16)	6; (2, 12)	0; (0, 1)	0; (0, 1)
VID.PA	7; (1, 18.5)	31; (14, 61)	28; (10, 68)	1; (0, 4)	2; (0, 10)
VID.PL	8; (1, 17.5)	31; (13, 83)	29; (13, 74)	1; (0, 6)	3; (0, 15)
SC_MT_TOT	15; (13, 17)	15; (11, 16)	13; (10, 15)	13; (10, 16)	16; (13, 17)
SC_FE_TOT	20; (15, 30.5)	17; (14, 26)	14; (11, 19)	17; (12, 24)	18; (14, 29)

Cluster 3: Engaged, but low performing (49, 17.1%)
Cluster 4: Disengaged (113, 39.4%)
Cluster 5: Engaged and well performing;
low in video-related activities (37, 13%)

Pre-midterm study approaches

Class 1 (35.4%)

Exercise-focused -> Exercise-focused -> Exercise-focused -> Exercise-focused, exhibiting 'guessing' behavior -> Disengaged

Class 2 (11.8%)

The Most Engaged -> The Most Engaged -> The Most Engaged | Engaged and high-performing ->
The most engaged, but not effective -> Highly engaged, exhibiting 'guessing' behavior | Disengaged

Pre-midterm study approaches

Class 3 (36.1%)

Engaged and high-performing -> The Most Engaged | Engaged and high-performing -> Engaged and high-performing -> Engaged and high-performing | Exercise-focused and high-performing -> Highly engaged, exhibiting 'guessing' behavior, focused on MCQs in lecture materials | Disengaged

Class 4 (16.7%)

Disengaged -> Disengaged -> Disengaged | Engaged and high-performing -> Disengaged -> Disengaged

Effects on grades

\	Q1	Median	Q3
c1	12.5	15	17.50
c2	10.0	13	15.75
c3	13.0	15	17.00
c4	10.0	11	15.00

Midterm exam

\	Q1	Median	Q3
c1	13.5	17	20.50
c2	11.0	15	20.75
c3	14.0	20	29.00
c4	13.0	17	23.00

Final exam

Differences (midterm) *: c1 > c4; c3 > c2; and c3 > c4

Differences (final)*: c3 > c2

*Kruskal-Wallis test followed by Mann-Whitney U test

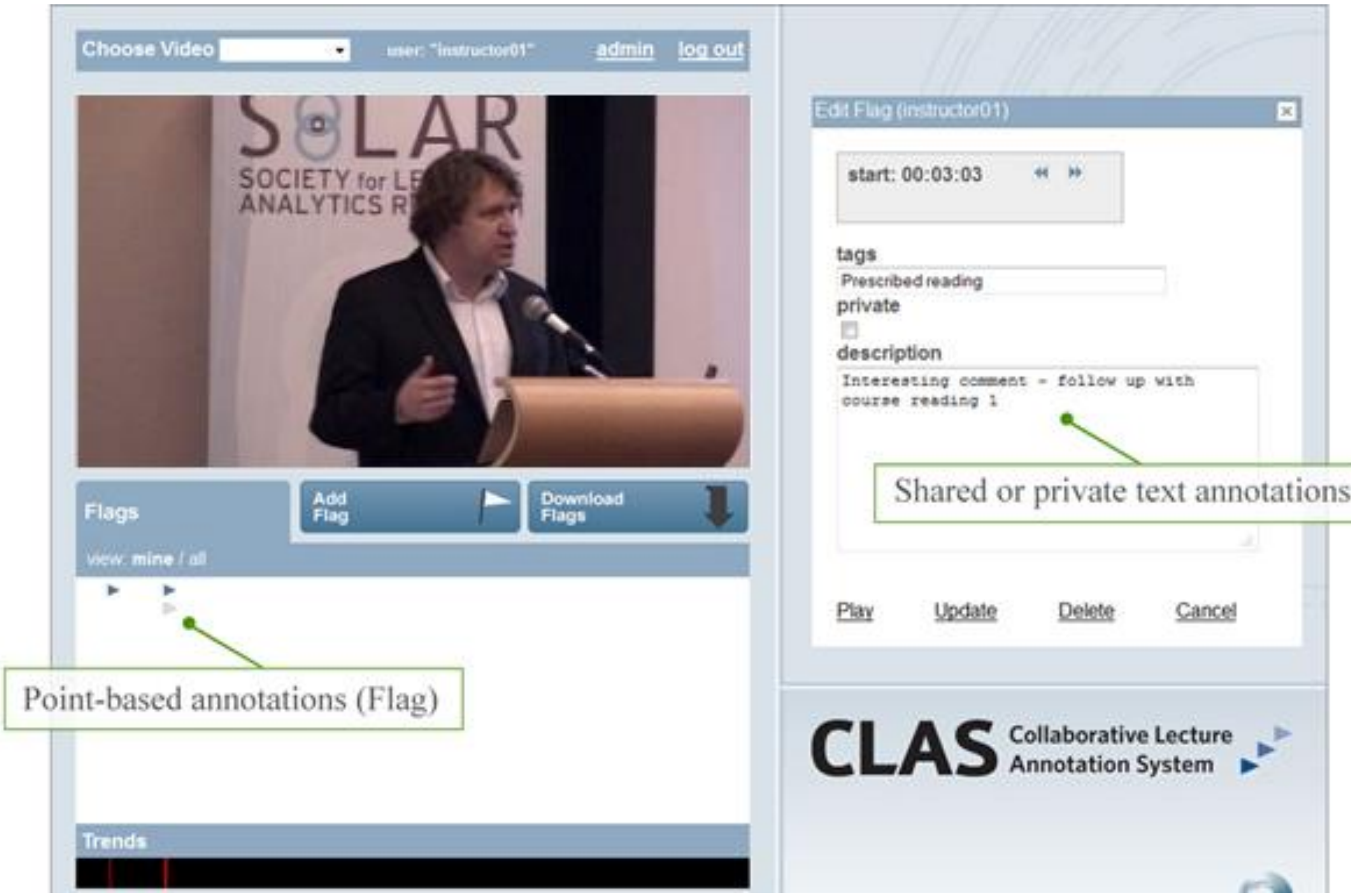
Transition matrix (probabilities) - HMM

From/To	Disengaged	Comprehensive use	Regular use	Strategic use
Disengaged	0.2426	0.2713	0.1183	0.3678
Comprehensive use	0.1310	0.4765	0.1970	0.1958
Regular use	0.2007	0.2380	0.2335	0.3279
Strategic use	0.1480	0.1267	0.0764	0.6489

LINKS OF LEARNING PROCESSES AND PRODUCTS

How are learning strategies associated with quality of learning products while using technology?

CLAS – Collaborative Lecture Annotation System



University of South Australia
THE UNIVERSITY OF NEW SOUTH WALES SYDNEY · AUSTRALIA
THE UNIVERSITY OF SYDNEY
UBC
THE UNIVERSITY OF EDINBURGH
Australian Government
Office for Learning and Teaching

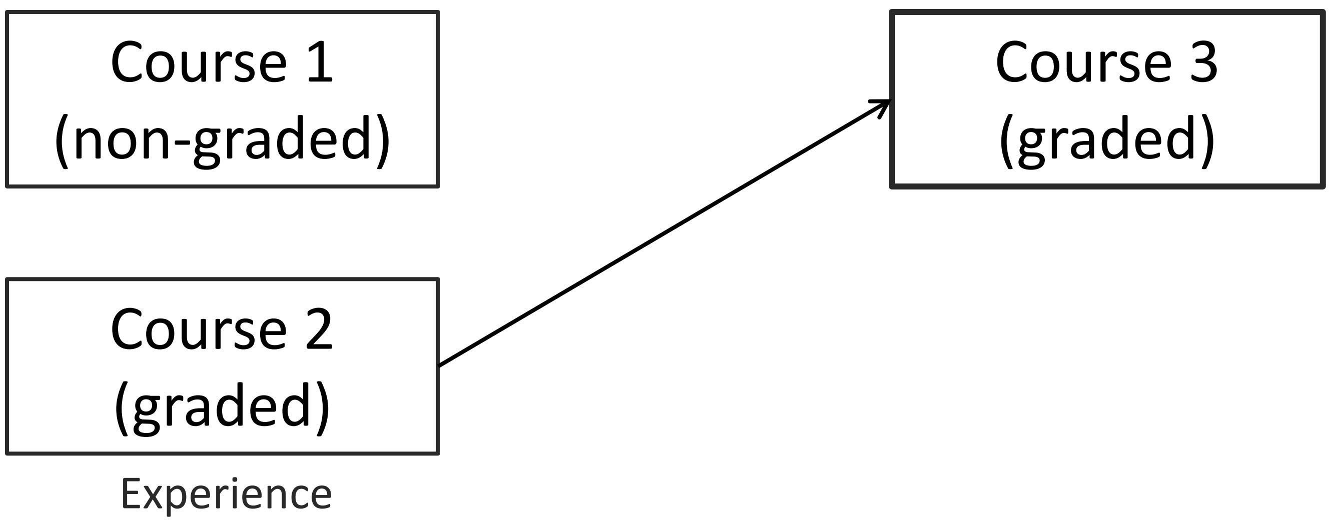
Self-reflections in video annotations

Course 1 (non-graded)

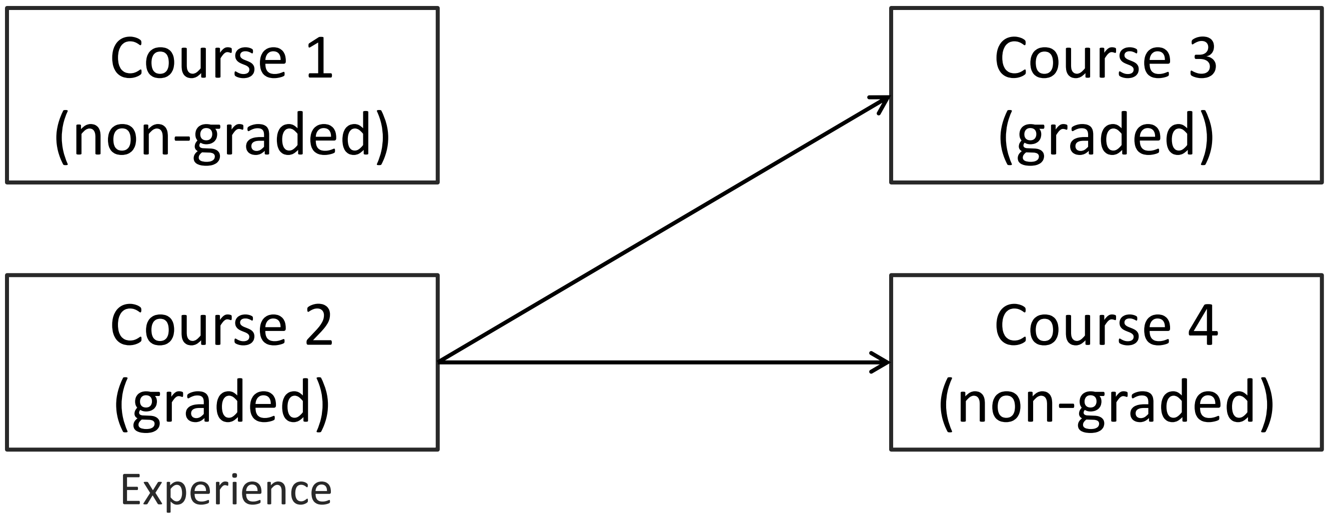
Course 2 (graded)

Experience

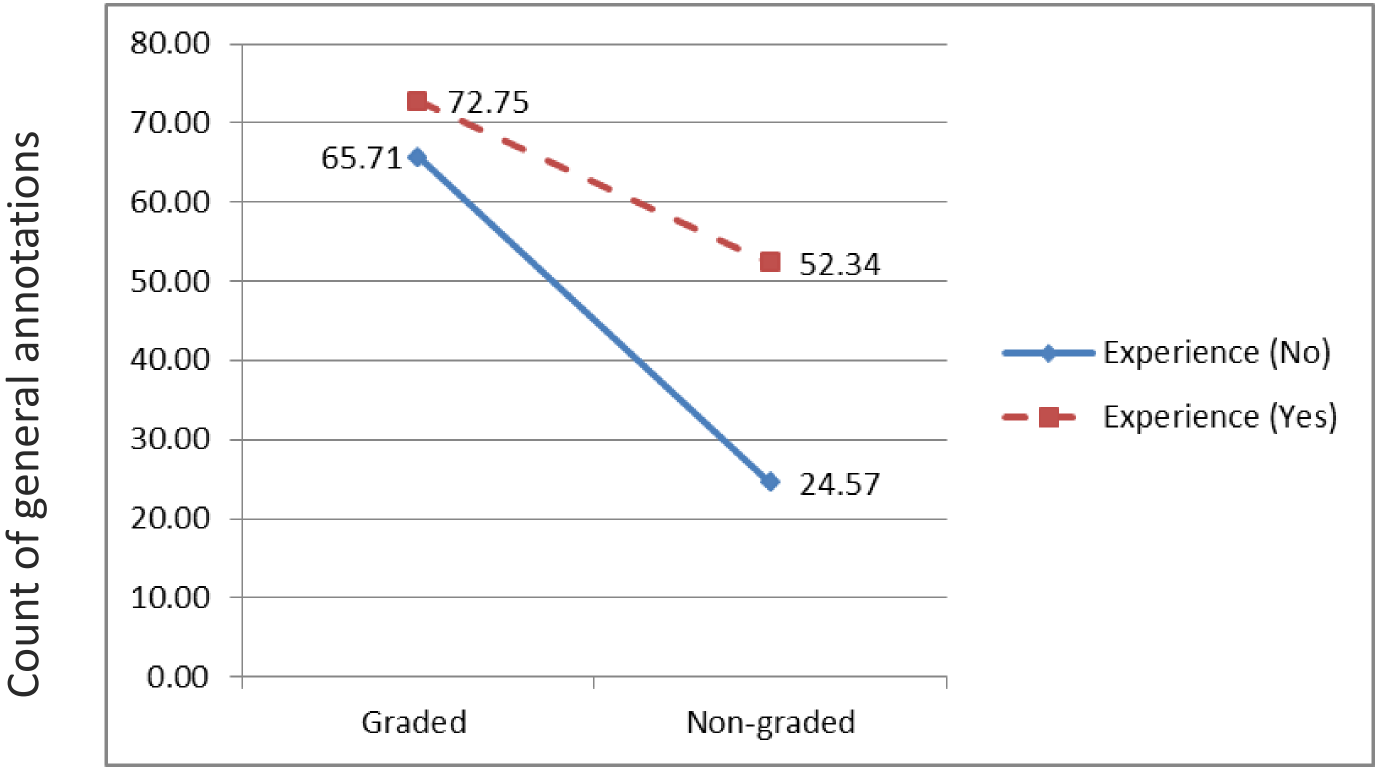
Self-reflections in video annotations



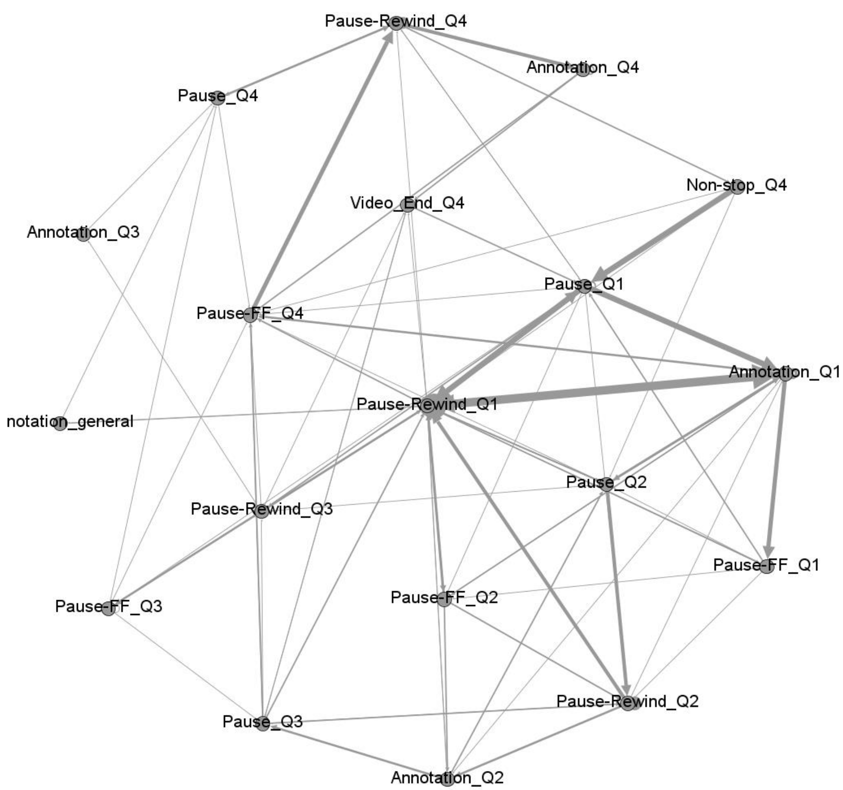
Self-reflections in video annotations



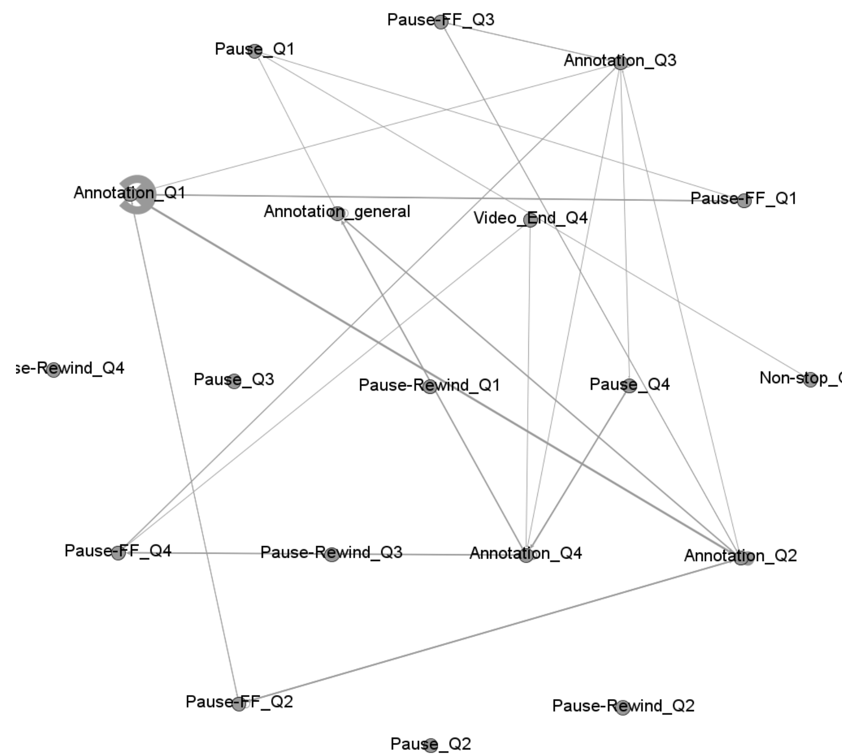
Annotation activity



Learning strategy -transition graphs-

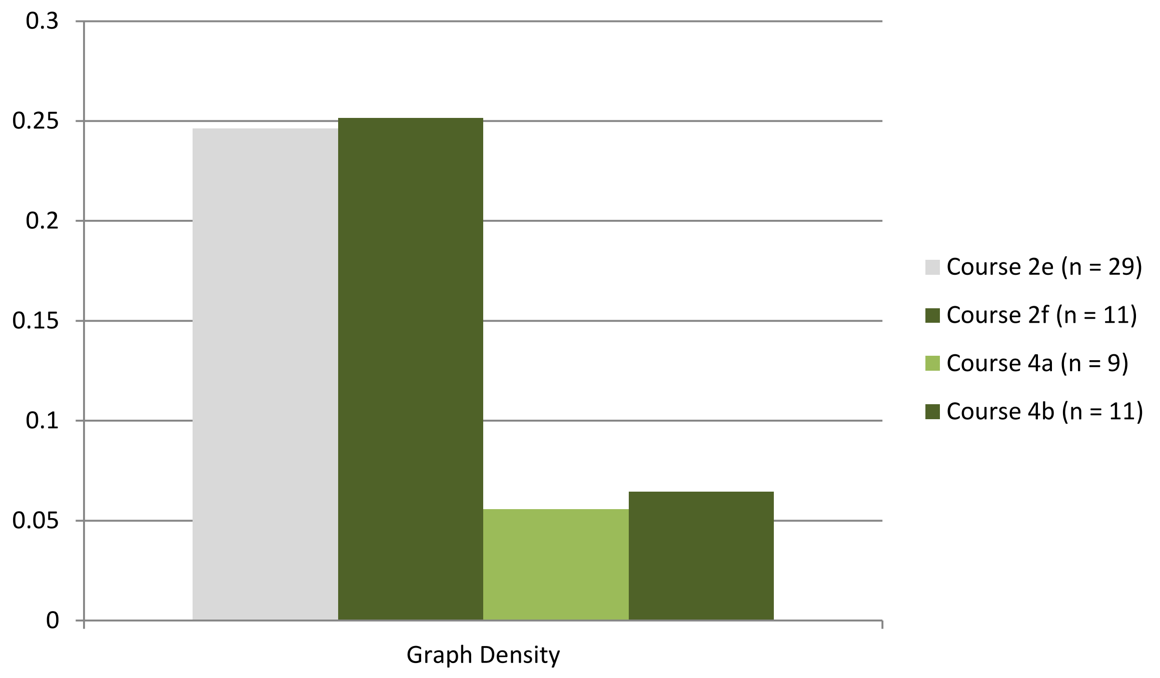


Student A
(course 2 – graded)



Student B
(course 4 – non-graded)

Transition graphs



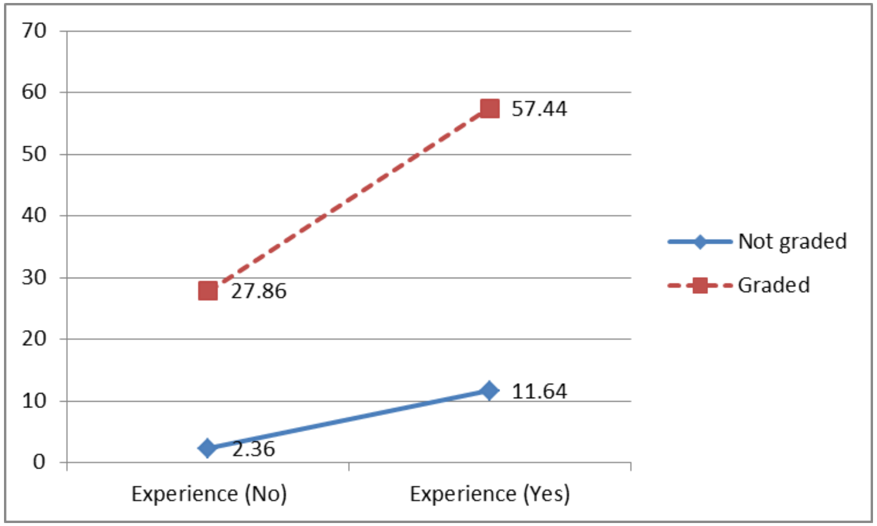
Reflection

Specificity associated with expertise development

Reflection specificity

Observation,
Motive/Effect, and Goal

Goal specificity



d) Goal

SCALING UP
QUALITATIVE METHODS

Community of Inquiry

Welcome | CoI Model | Papers | News | Contact

Welcome

This site documents the work completed during a Canadian Social Sciences and Humanities research funded project entitled "A Study of the Characteristics and Qualities of Text-Based Computer Conferencing for Educational Purposes". This project ran from 1997 to 2001. The theory, methodology and instruments developed during this project are described in the papers published in peer reviewed journals and copied at this site.

The work of this project has resulted in a variety of researchers replicating and further developing the tools and techniques that we developed. We invite anyone who uses this content to contribute their own papers, references, and links in the related sections. As well, feel free to share experiences, concerns or questions in the weblog. The purpose of this project is to support a personally meaningful and educationally worthwhile learning experience. Central to the study introduced here is the model of a community of inquiry that constitutes three elements essential to an educational experience: Cognitive Presence, Social Presence and Teaching Presence.

Cognitive presence

Triggering event

Exploration

Integration

Resolution

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical Thinking and Computer Conferencing: A Model and Tool to Assess Cognitive Presence. *American Journal of Distance Education, 15*(1)*, 7-23*.

Manual analysis is labor intensive

Cognitive presence classifier

	Predicted				
Actual	Other	Triggering	Explorat.	Integrat.	Resolut.
Other	**79**	2	2	2	2
Triggering	5	**67**	9	6	0
Exploration	9	15	**35**	27	1
Integration	2	2	23	**44**	16
Resolution	0	0	4	2	**81**

Random forest
Features: Named entities, LIWC features, LSA features Coh-Metrix features, and contextual
Cohen's $\kappa = 0.65$

Kovanović, V., Joksimović, S., Waters, Z., Gašević, D., Kitto, K., Hatala, M., Siemens, G. (in press). Towards Automated Content Analysis of Discussion Transcripts: A Cognitive Presence Case. In *Proceedings of the 5th International Conference on Learning Analytics & Knowledge (LAK 2016)*, Edinburgh, Scotland, UK, 2016.

Cognitive presence classifier

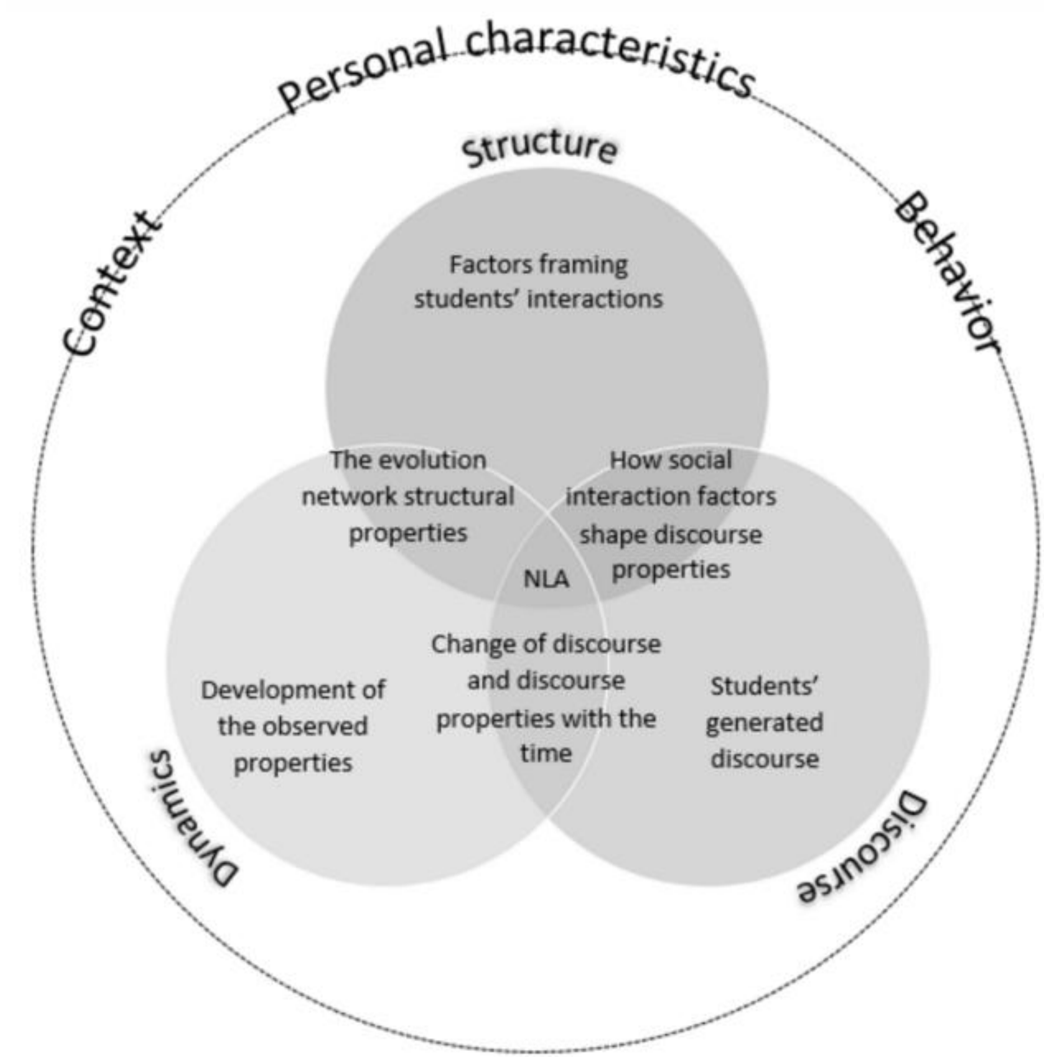
#	Variable	Description	MDG*	Cognitive presence phase				
				Other	**Triggering**	**Exploration**	**Integration**	**Resolution**
1	cm.DESWC	Number of words	32.91	55.41 (61.06)	80.91 (41.56)	117.71 (67.23)	183.30 (102.94)	280.68 (189.62)
2	ner.entity.cnt	Number of named entities	26.41	13.44 (15.36)	21.67 (10.55)	28.84 (16.93)	44.75 (24.85)	64.18 (32.54)
3	cm.LDTTRa	Lexical diversity, all words	21.98	0.85 (0.12)	0.77 (0.09)	0.71 (0.10)	0.65 (0.09)	0.58 (0.09)
4	message.depth	Position within discussion	19.09	2.39 (1.13)	1.00 (0.90)	1.84 (0.97)	1.87 (0.94)	2.00 (0.68)
5	cm.LDTTRc	Lexical diversity, content words	17.12	0.95 (0.06)	0.90 (0.06)	0.86 (0.08)	0.82 (0.07)	0.78 (0.07)
6	cm.LSAGN	Avg. givenness of each sentence	16.63	0.10 (0.07)	0.14 (0.06)	0.18 (0.07)	0.21 (0.06)	0.24 (0.06)
7	liwc.QMark	Number of question marks	16.59	0.27 (0.85)	1.84 (1.63)	0.92 (1.26)	0.58 (0.82)	0.38 (0.55)
8	message.sim.prev	Similarity with previous message	16.41	0.20 (0.17)	0.06 (0.13)	0.22 (0.21)	0.30 (0.24)	0.39 (0.19)
9	cm.LDVOCD	Lexical diversity, VOCD	15.43	12.92 (33.93)	28.99 (50.61)	53.57 (54.68)	83.47 (43.00)	97.16 (28.95)
10	liwc.money	Number of money-related words	14.38	0.21 (0.69)	0.32 (0.74)	0.32 (0.75)	0.65 (1.12)	0.99 (1.04)
11	cm.DESPL	Avg. number of paragraphs sent.	12.47	4.26 (2.98)	6.37 (2.76)	7.49 (4.11)	10.17 (5.64)	14.05 (8.88)
12	message.sim.next	Similarity with next message	11.74	0.08 (0.14)	0.34 (0.40)	0.20 (0.22)	0.22 (0.24)	0.22 (0.23)
13	message.reply.cnt	Number of replies	11.67	0.42 (0.67)	1.44 (1.89)	0.82 (1.70)	1.10 (2.66)	0.84 (1.24)
14	cm.DESSC	Sentence count	11.67	4.28 (3.17)	6.36 (2.75)	7.49 (4.11)	10.17 (5.64)	14.29 (10.15)
15	lsa.similarity	Avg. LSA sim. between sentences	9.69	0.29 (0.27)	0.47 (0.23)	0.54 (0.23)	0.62 (0.20)	0.67 (0.17)
16	cm.DESSL	Avg. sentence length	9.60	11.88 (6.82)	13.62 (5.85)	16.69 (6.54)	19.36 (8.39)	21.73 (8.61)
17	cm.DESWLsyd	SD of word syllables count	8.92	0.98 (0.69)	1.33 (0.70)	0.98 (0.18)	0.97 (0.14)	0.97 (0.11)
18	liwc.i	Number of FPS* pronouns	8.84	4.33 (3.53)	2.82 (2.06)	2.37 (1.94)	2.51 (1.65)	2.19 (1.23)
19	cm.RDFKGL	Flesch-Kincaid Grade Level	8.29	7.68 (4.28)	10.30 (3.50)	10.19 (3.11)	11.13 (3.46)	11.99 (3.37)
20	cm.SMCAUSwn	WordNet overlap between verbs	8.14	0.38 (0.25)	0.48 (0.20)	0.51 (0.13)	0.50 (0.10)	0.47 (0.06)

MDG - Mean decrease Gini impurity index, **FPS** - first person singular

Kovanović, V., Joksimović, S., Waters, Z., Gašević, D., Kitto, K., Hatala, M., Siemens, G. (in press). Towards Automated Content Analysis of Discussion Transcripts: A Cognitive Presence Case. In *Proceedings of the 5th International Conference on Learning Analytics & Knowledge (LAK 2016)*, Edinburgh, Scotland, UK, 2016.

MIXING ANALYTICS METHODS

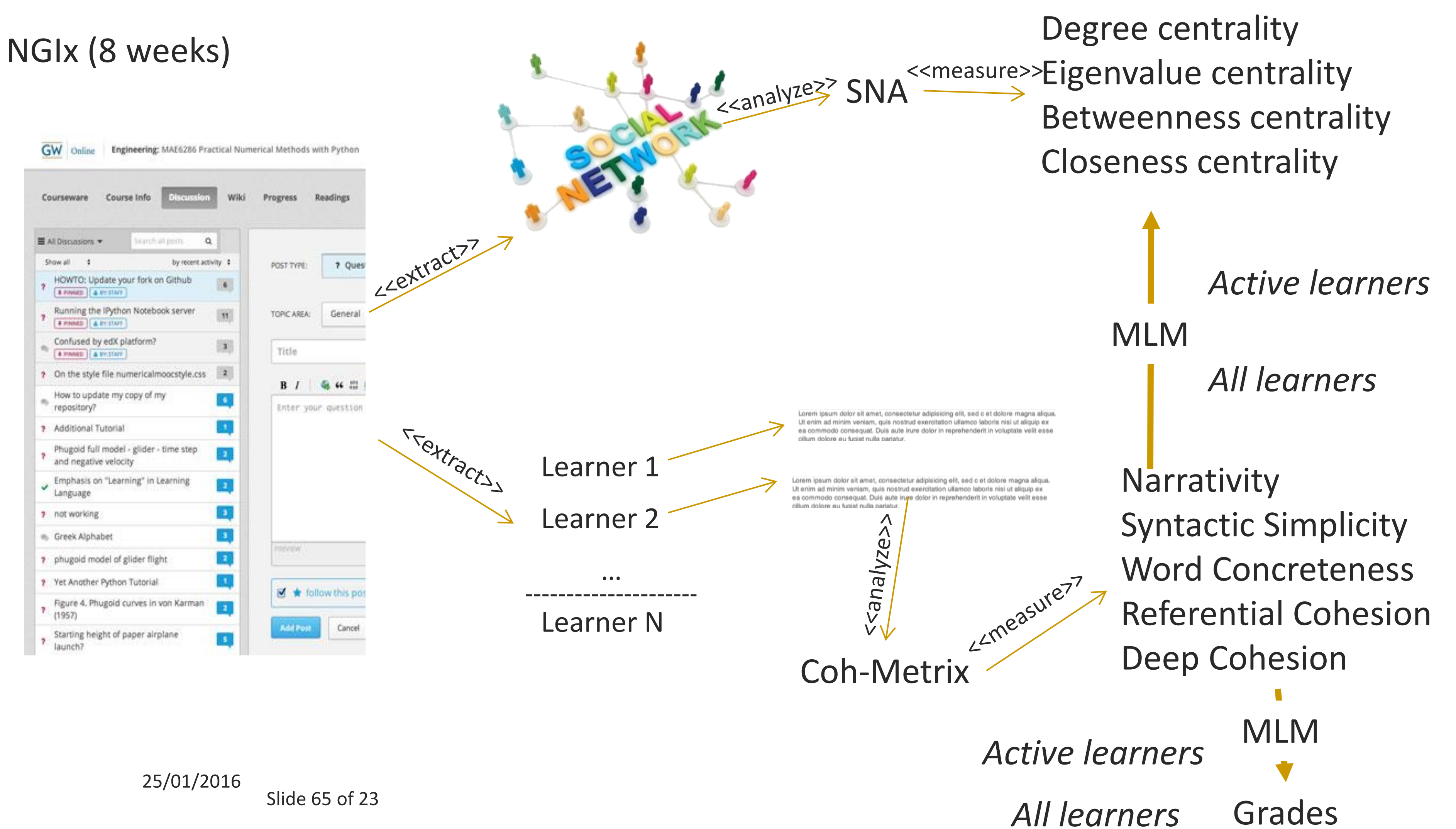
Network learning analytics



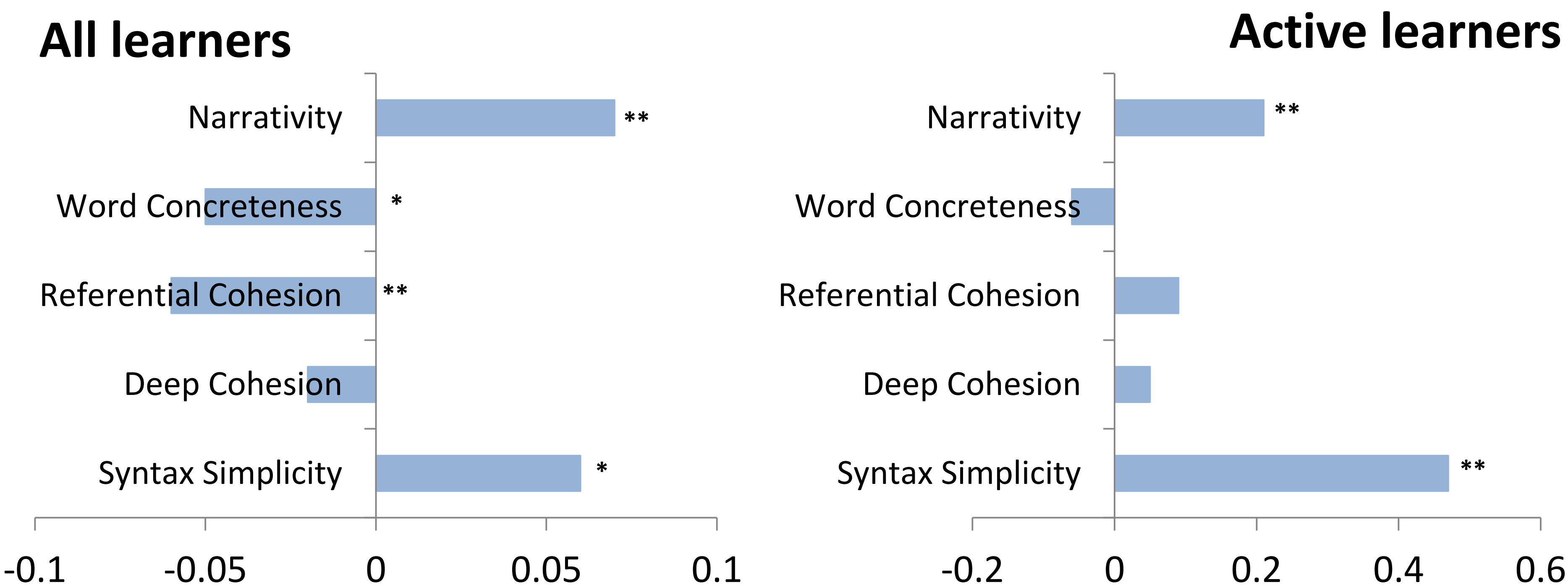
How does language shape network centrality and performance?

Joksimović, S., Dowell, N., Skrypnyk, O., Kovanović, V., Gašević, D., Dawson, S., Graesser, A. C. (under review). Exploring the Accumulation of Social Capital in cMOOC Through Language and Discourse," *International Review of Research in Online and Distance Learning.*

xMOOC study approach

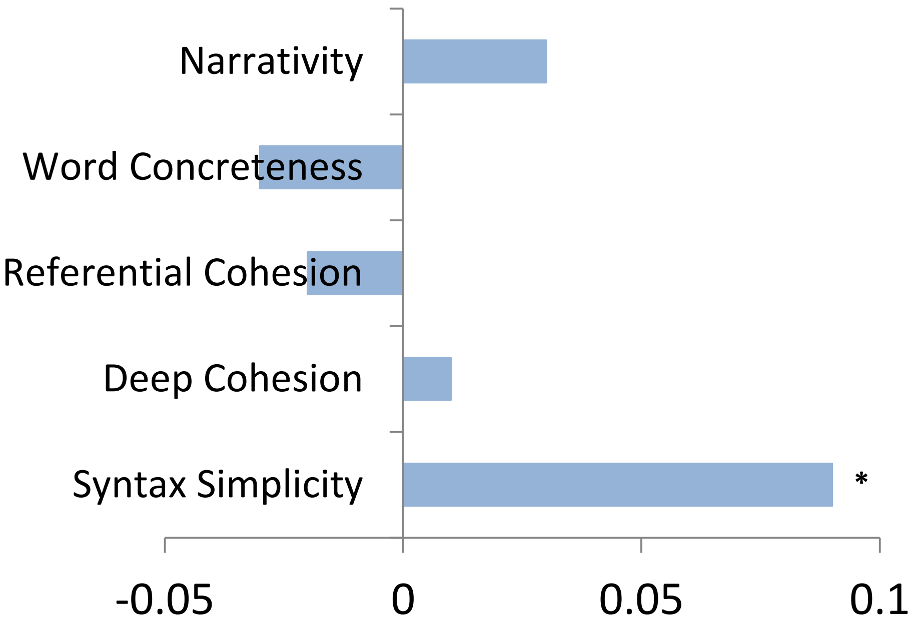


Degree centrality

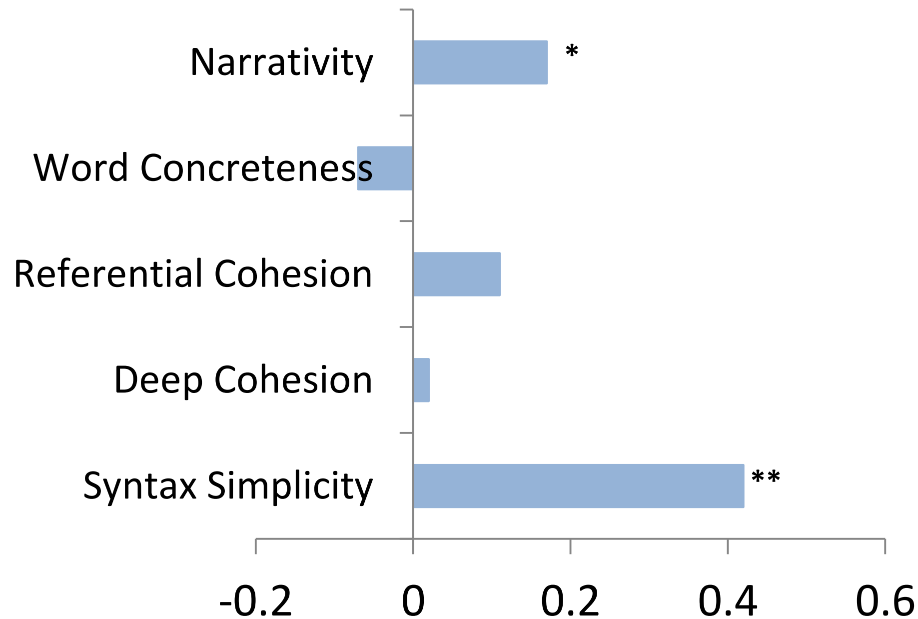


Betweenness centrality

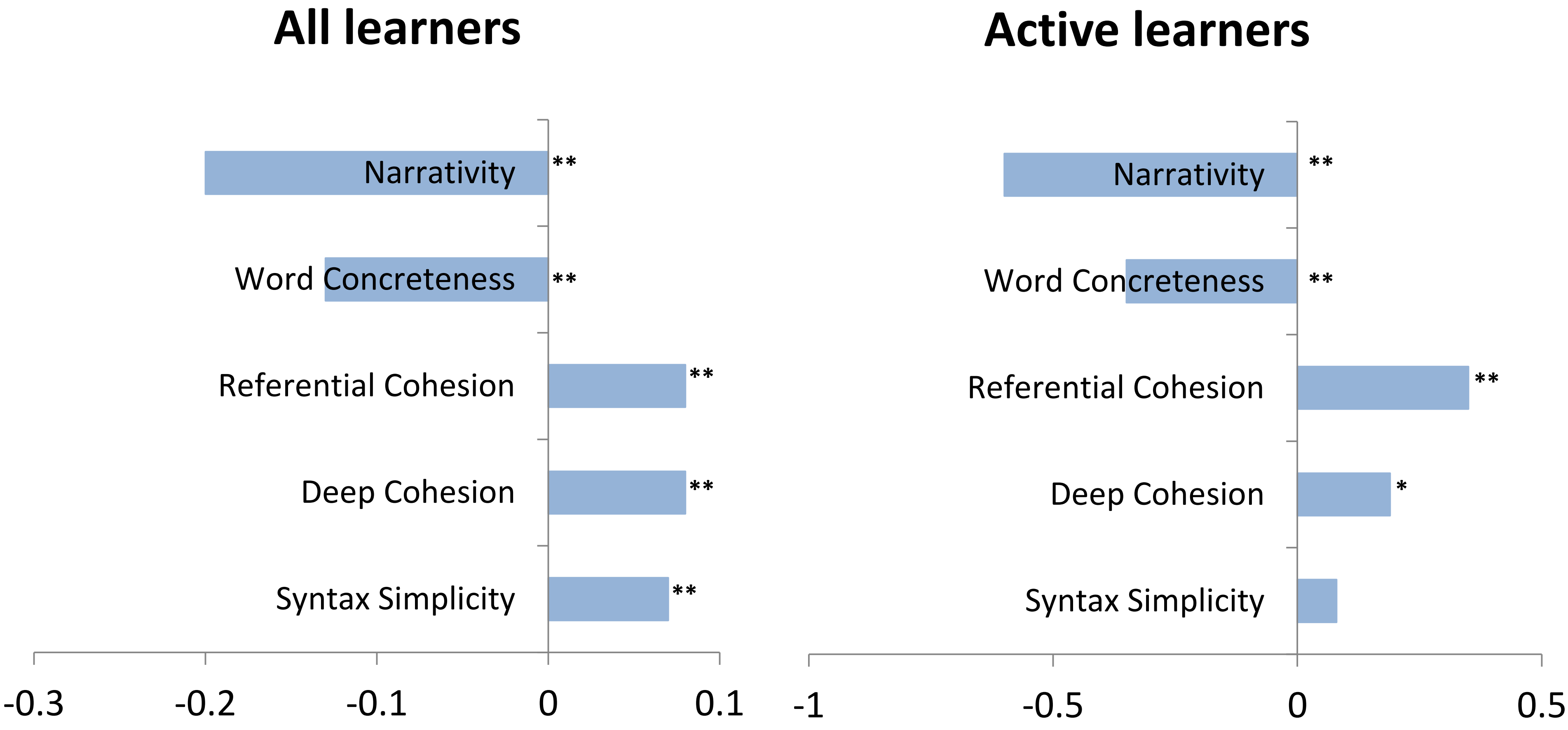
All learners



Active learners



xMOOC *Performance models*



NEW SOURCES OF DATA

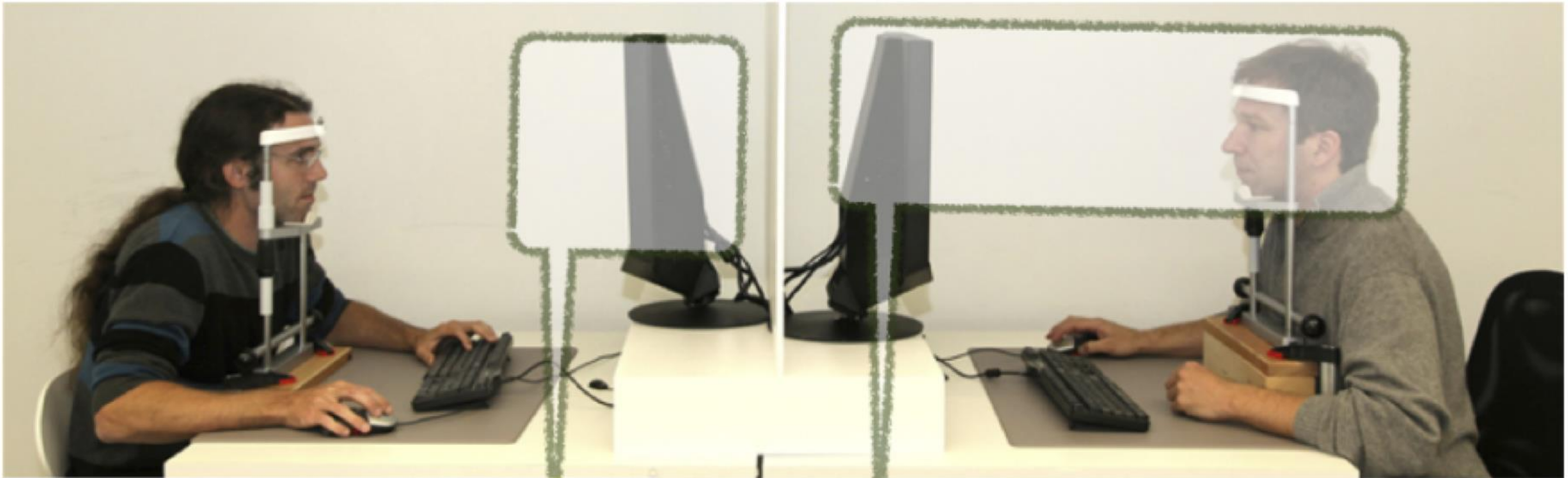
Physiological measurement and wearables

Eye gazing to track the sync of students with video lectures in MOOCs

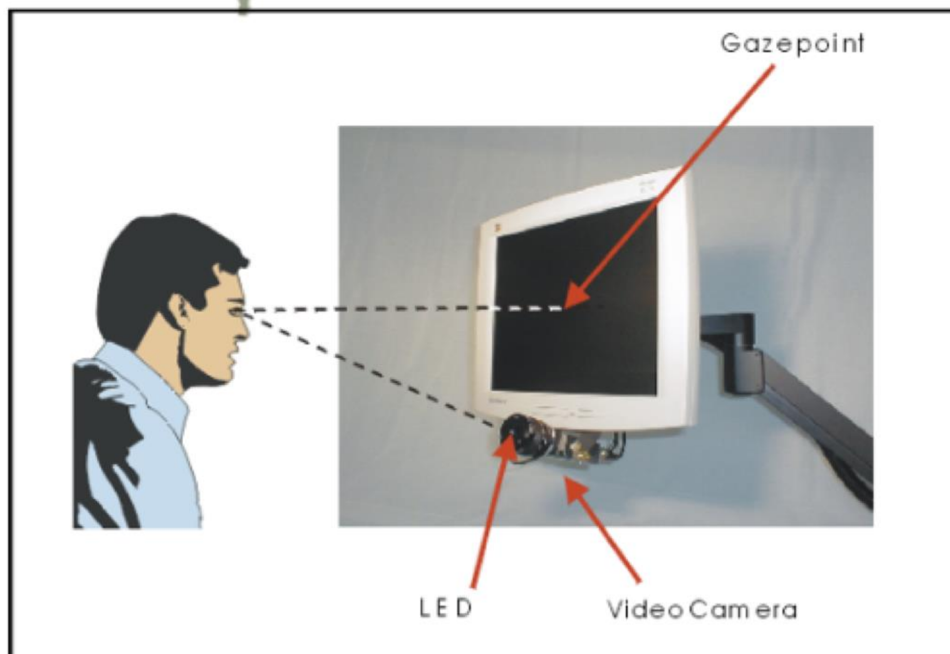
Sharma, K., Caballero, D., Verma, H., Jermann, P., & Dillenbourg, P. (2015). Looking AT versus Looking THROUGH: A Dual Eye-tracking study in MOOC Context. In Proceedings of *Computer Supported Collaborative Learning*.

Dual eye gazing to track student collaboration success

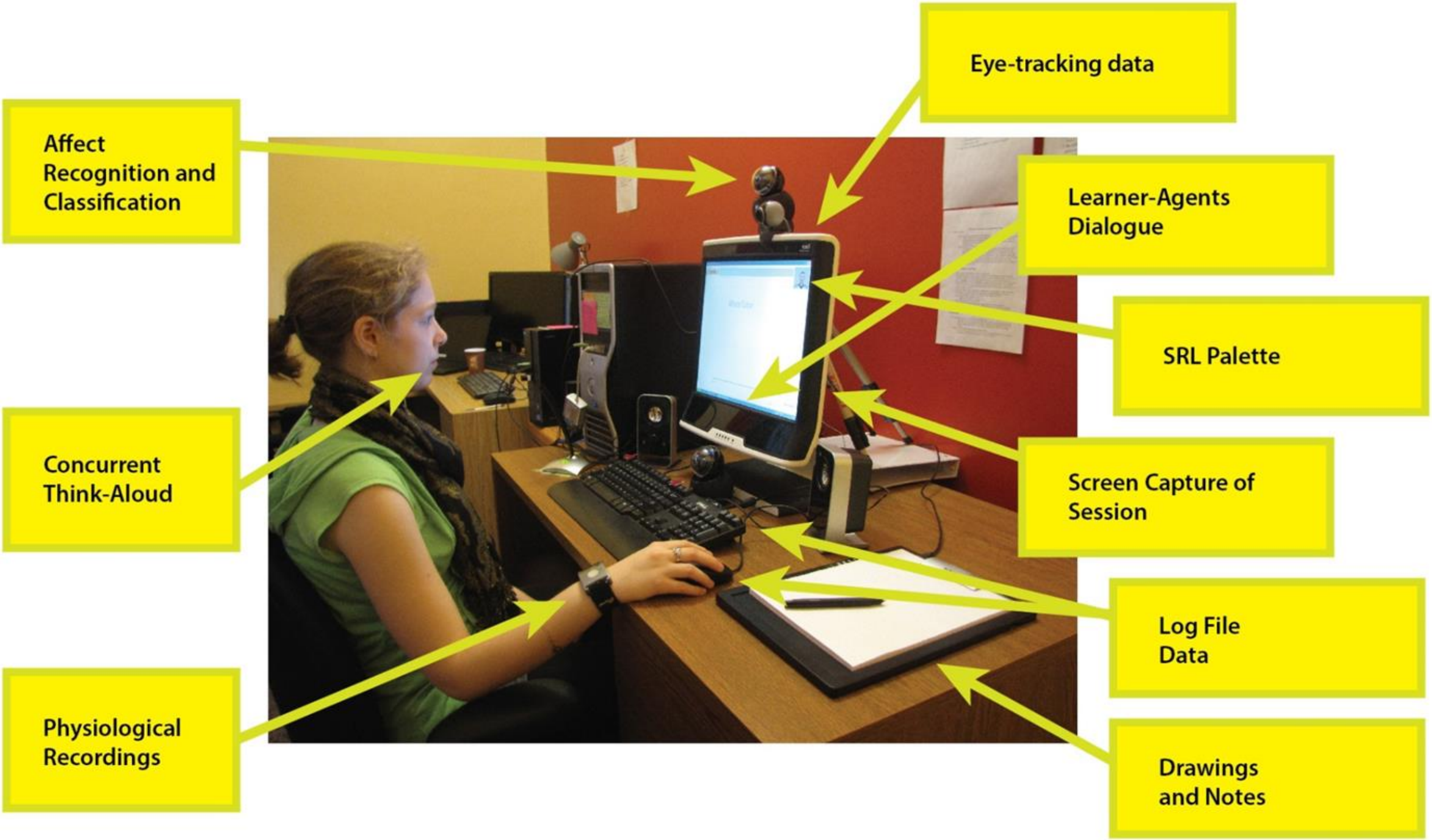
Sharma, K., Caballero, D., Verma, H., Jermann, P., & Dillenbourg, P. (2015). Looking AT versus Looking THROUGH: A Dual Eye-tracking study in MOOC Context. In Proceedings of *Computer Supported Collaborative Learning*.



```
for (c = 1; c < size; c++)
{
  if (array[c] > maximum)
  {
     maximum  = array[c];
     location = c+1;
  }
}
```

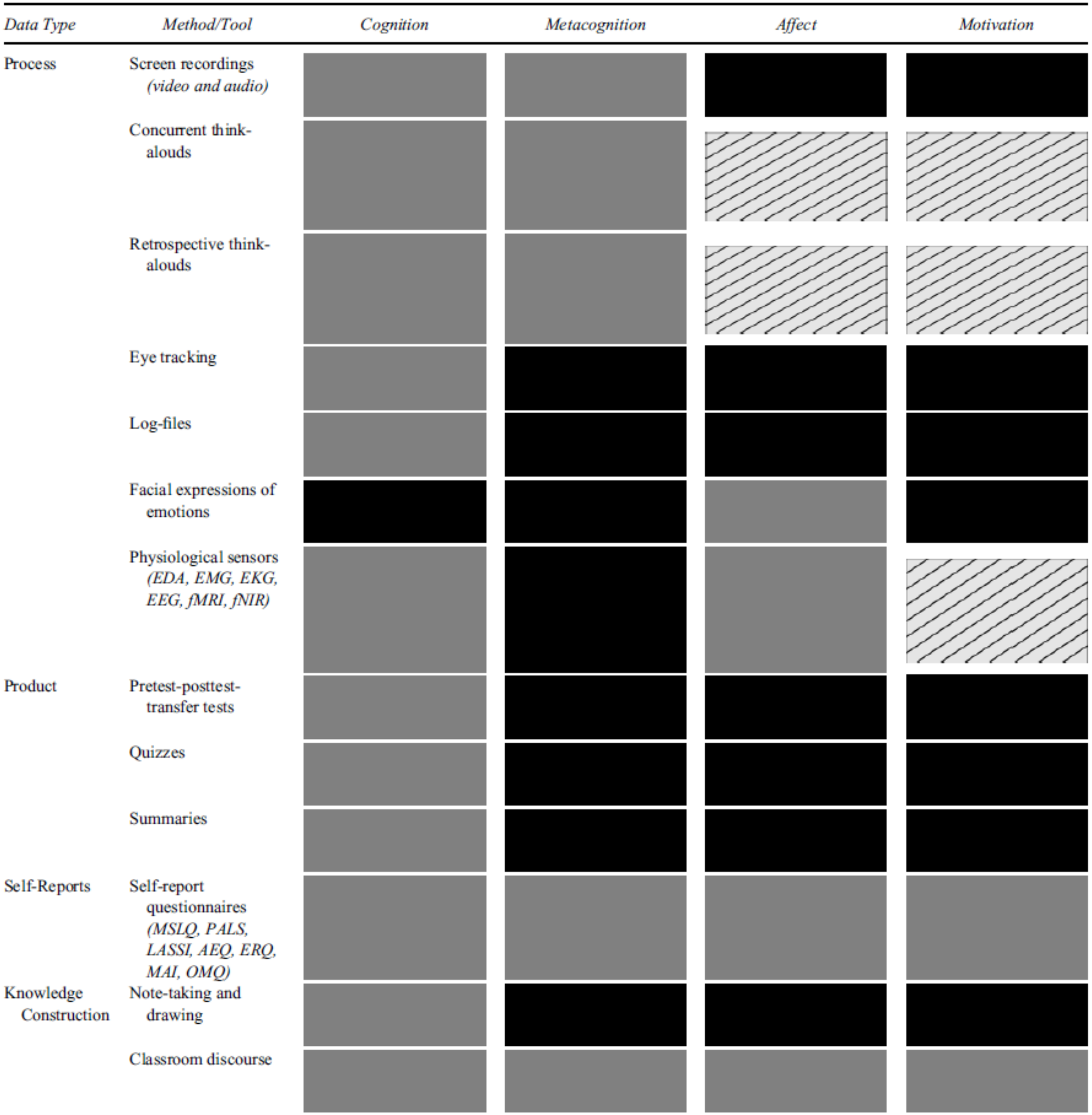


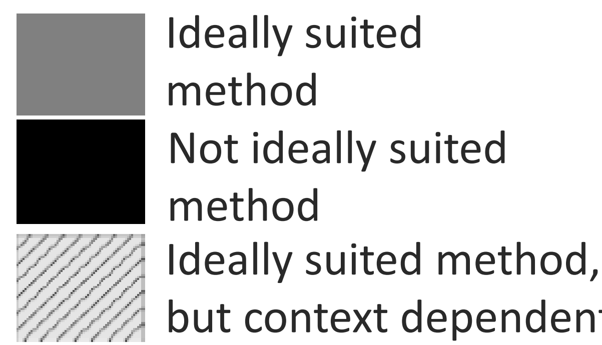
Sharma, K., Jermann, P., Nüssli, M. A., & Dillenbourg, P. (2013). Understanding collaborative program comprehension: Interlacing gaze and dialogues. In *Proceedings of Computer Supported Collaborative Learning (CSCL 2013)*.



Multi-modality Lab at North Carolina State University (courtesy of Roger Azevedo)

Capturing and measurement of engagement-related processes





Azevedo, R. (2015). Defining and measuring engagement and learning in science: Conceptual, theoretical, methodological, and analytical issues. *Educational Psychologist*, *50*(1), 84-94.

CONCLUSION

Process nature of learning
- beyond coding and counting -

van der Aalst, W. (2012). Process mining: Overview and opportunities. *ACM Transactions on Management Information Systems (TMIS)*, *3*(2), 7.

Scaling up qualitative analysis

Approaches to mixing data sources and analysis methods

More granular trace data for real-time feedback

Yudelson, M. V., Koedinger, K. R., & Gordon, G. J. (2013, January). Individualized Bayesian knowledge tracing models. In *Artificial Intelligence in Education* (pp. 171-180). Springer Berlin Heidelberg.

Can we make more dynamic and self-adaptive models?

Better instrumentation and measurement needed

Design principles and effects of analytics-based feedback

Ethics and privacy in learning analytics

Learning Analytics Summer Institute
University of Michigan
June 27-29 2016

SOLAR
SOCIETY for LEARNING ANALYTICS RESEARCH

Many thanks!

LAK 16

The 6th International
Learning Analytics & Knowledge Conference
University of Edinburgh, Edinburgh, UK, April 25-29, 2016

SOLAR
SOCIETY for LEARNING ANALYTICS RESEARCH