

CS-411 : Digital Education

Chapter 11 From learning analytics to classroom analytics

What if the learning activity
does not work well for Pierre ?

The learner adapts

If you don't understand the video,
play it again but slower

The teacher adapts

If students don't understand the lecture,
re-explain

The system adapts

If students don't master the skill, select a
more effective learning activity

Adaptive instruction

Personalized instruction

Individualized instruction

~~Recommender system~~



Learner Modelling

$$5^2 = ?$$

Learner Modelling

From the learner's **behaviour**, infer his/her **knowledge state**



$$5^2 = 25 \rightarrow$$

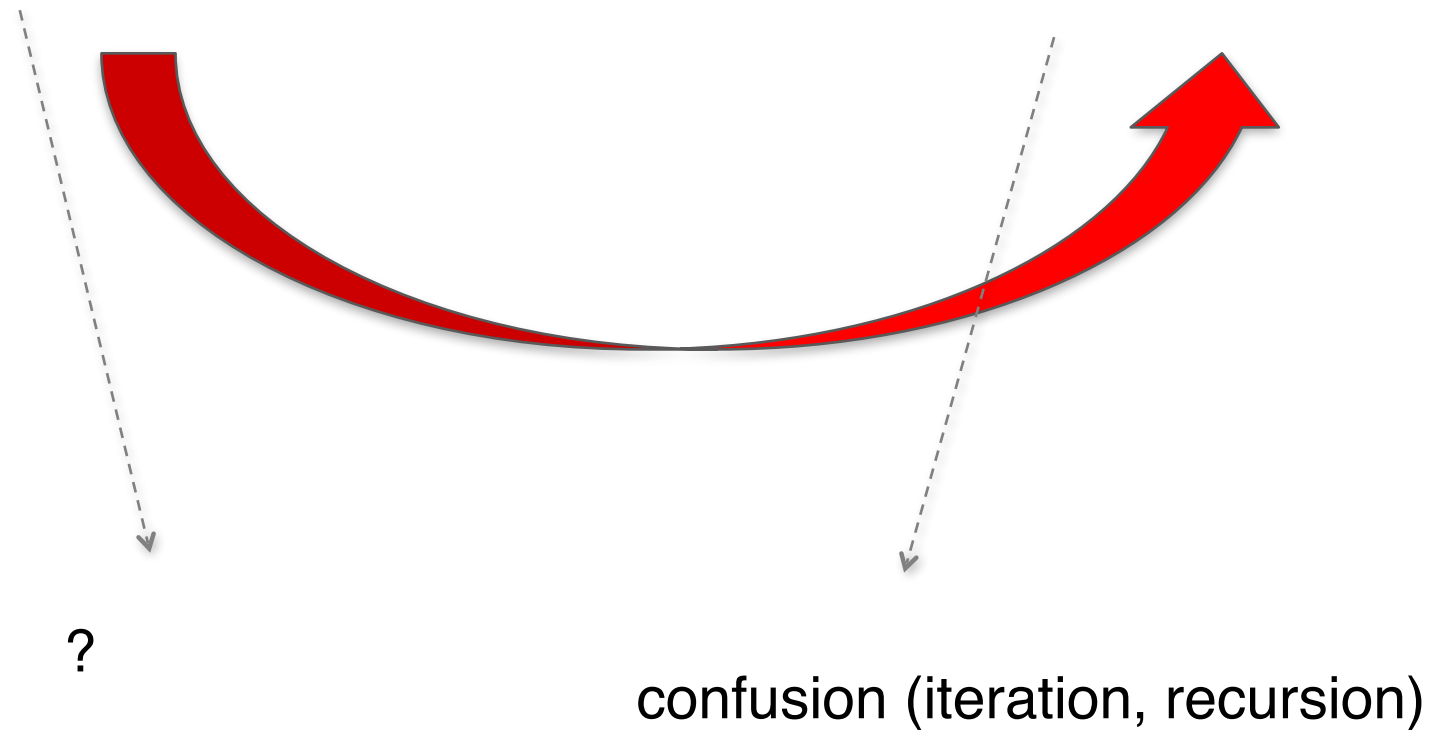
knows X^2

$$5^2 \neq 25 \rightarrow$$

doesn't know X^2

Learner Modelling

From the learner's **behaviour**, infer his/her **knowledge state**



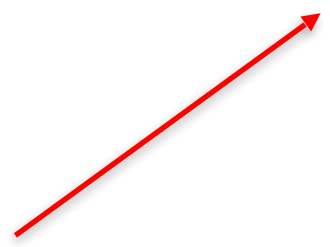
Learner Modelling

$5^2 =$	☐	52	p (knows X^2) = 1 ?
	☒	25	
	☐	15	
	☐	10	
	☐	7	

Cognitive Diagnosis

$5^2 = ??$	Knowledge States										
Behavior (Answer)	$5^2 = 25$	$5^n = \dots$	$n^2 = n \cdot N$	$x^n = x \cdot x \dots$	$x^n = x \cdot x$ but bad mult.	$x^n = x \cdot n$	$x^n = x + n$	$x^n = ???$	Sum	Entropy	Normalized entropy
25	0.10	0.20	0.30	0.40	0.00	0.00	0.00	0.00	1	1.89	0.63
35	0.00	0.00	0.00	0.00	0.40	0.10	0.00	0.50	1	1.41	0.47
10	0.00	0.00	0.00	0.00	0.00	0.79	0.00	0.20	1	0.79	0.26
27	0.00	0.00	0.00	0.00	0.40	0.00	0.00	0.59	1	1.03	0.34
7	0.00	0.00	0.00	0.00	0.00	0.00	0.59	0.40	1	1.03	0.34
											0.41

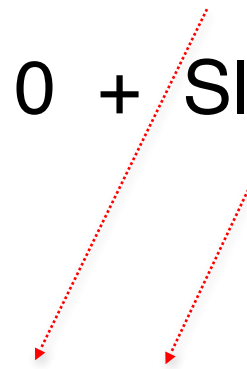
Diagnosis Power
(The lower the better)



Learner Modelling

From the learner's **behaviour**, infer his/her **knowledge state**

$$p(\text{state} = \text{knows} \mid \text{correct-answer}) = 1 - \text{Guess}$$

$$p(\text{state} = \text{knows} \mid \text{incorrect-answer}) = 0 + \text{Slip}$$


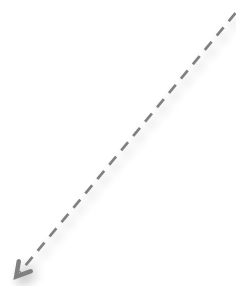
Factors that depend upon the response modality

Knowledge States				
$n \in \mathbb{N}$	$x^n = x.x....$	$x^n = x \cdot x$ but bad mult.	$x^n = x.n$	$x^n = x +$
0.30	0.40	0.00	0.00	0.00
0.00	0.00	0.40	0.10	0.00
0.00	0.00	0.00	0.79	0.00
0.00	0.00	0.40	0.00	0.00
0.00	0.00	0.00	0.00	0.50

[illegible]

In a MOOC ?

From the learner's **behaviours**, infer his/her **knowledge state**



$b(s)$ = watch video with many pauses

$b(s)$ = select correct definition of SD in a quiz with 5 possible definitions

$b(s)$ = post a message "There is a mistake on the slide" (and there is one indeed)



$X(S) = \{\text{lost, active, fine, brilliant}\}$

$x(s) = [.15 \ .40 \ .30 \ .15]$

$x(s) = [.05 \ .15 \ .25 \ .55]$

$x(s) = [.01 \ .02 \ .02 \ .95]$

$H_0 = 0.94$

$H_0 = .80$

$H_0 = .18$

Normalized entropy of the diagnosis vector

Which question has the highest diagnosis power ?

Question 1

The standard deviation of a distribution is theof the sum of from the mean

Question 2

Remove two numbers from this distribution to minimize it's standard deviation : [1 3 3 5 9 9 9 10 11 18 19 25 29]

$$x(s) = [.15 \ .40 \ .30 \ .15]$$

$$H(X) = - \sum_i P(x_i) \log_b P(x_i)$$

$$H_0 = 0.94$$

The uncertainty of the diagnosis can be estimated by Shannon's entropy applied to the vector of probabilities for the different states.

Since this value depends upon the number of states, we normalize it on a 0->1 scale by dividing it by the maximal entropy which is $\log_2$ of the number of states

The **diagnosis power** of a question is directly proportional to the normalized entropy of the diagnosis vector

Decrease uncertainty by collecting multiple answers

[illegible]

Decrease uncertainty by collecting multiple answers

[illegible]

How does the teacher/system
chooses the next question ?

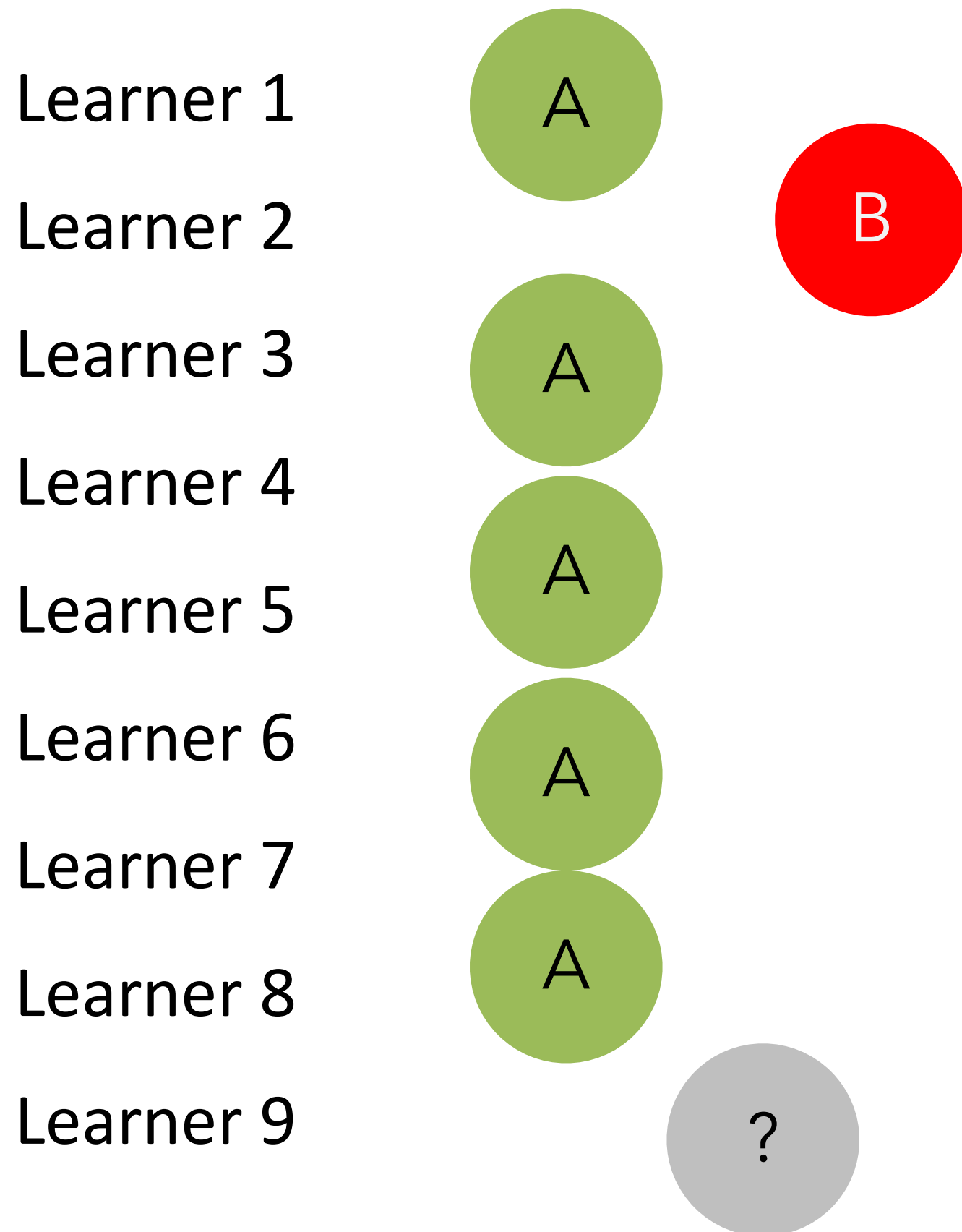
- Because it will maximize the learning gain of the learner ?

Exploitation

- Because it will maximize the system knowledge about the learner ?

Exploration

Exploration Exploitation Tradeoff



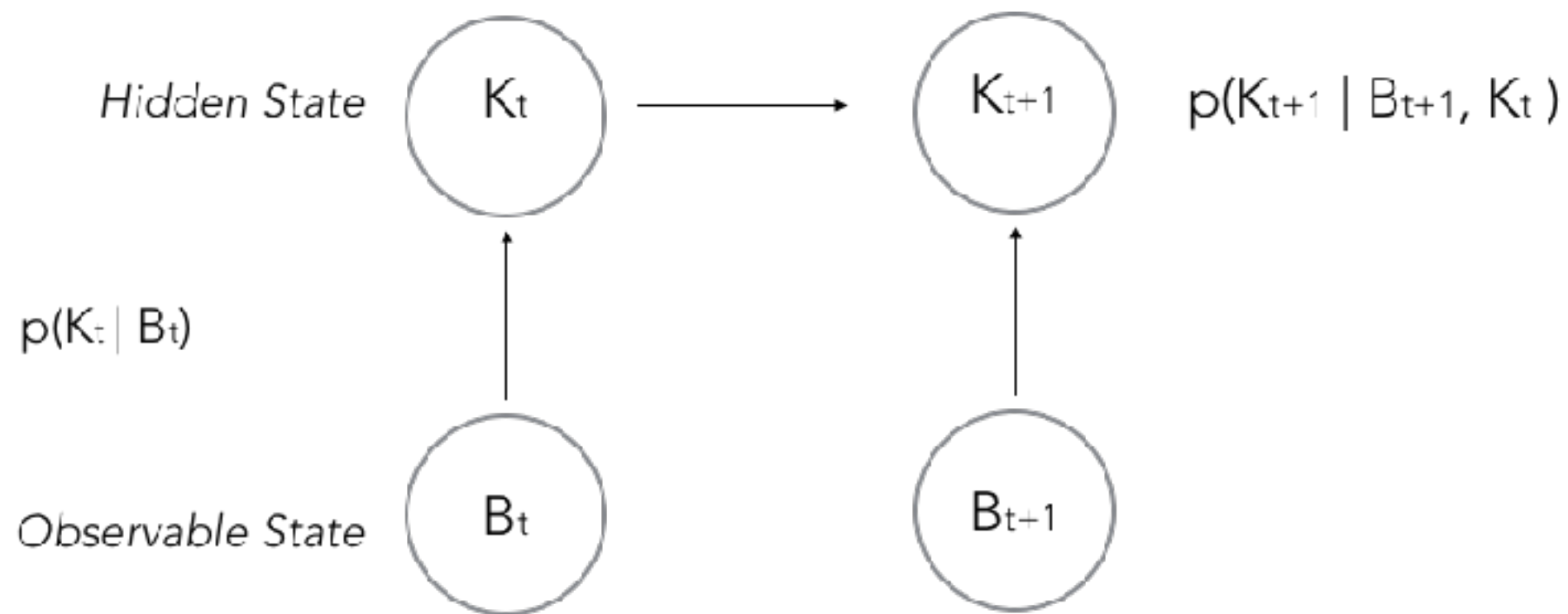
Learner Modelling

From the learner's behaviour, **infer** his/her knowledge state

+

From the learner's previous state, **predict** his/her knowledge state

From the learner's behaviour and **his previous state**, infer his/her knowledge state



Bayesian Knowledge Tracing

CS-421

Machine learning for behavioral data

Käser Tanja

Cursus	Sem.	Type
Cybersecurity	MA2, MA4	Opt.
Data Science	MA2, MA4	Opt.
Informatique	MA2, MA4	Opt.
SC master EPFL	MA2, MA4	Opt.

Language	English
Credits	4
Session	Summer
Semester	Spring
Exam	Written
Workload	120h
Weeks	14
Hours	4 weekly
Lecture	2 weekly
Project	2 weekly
Number of positions	

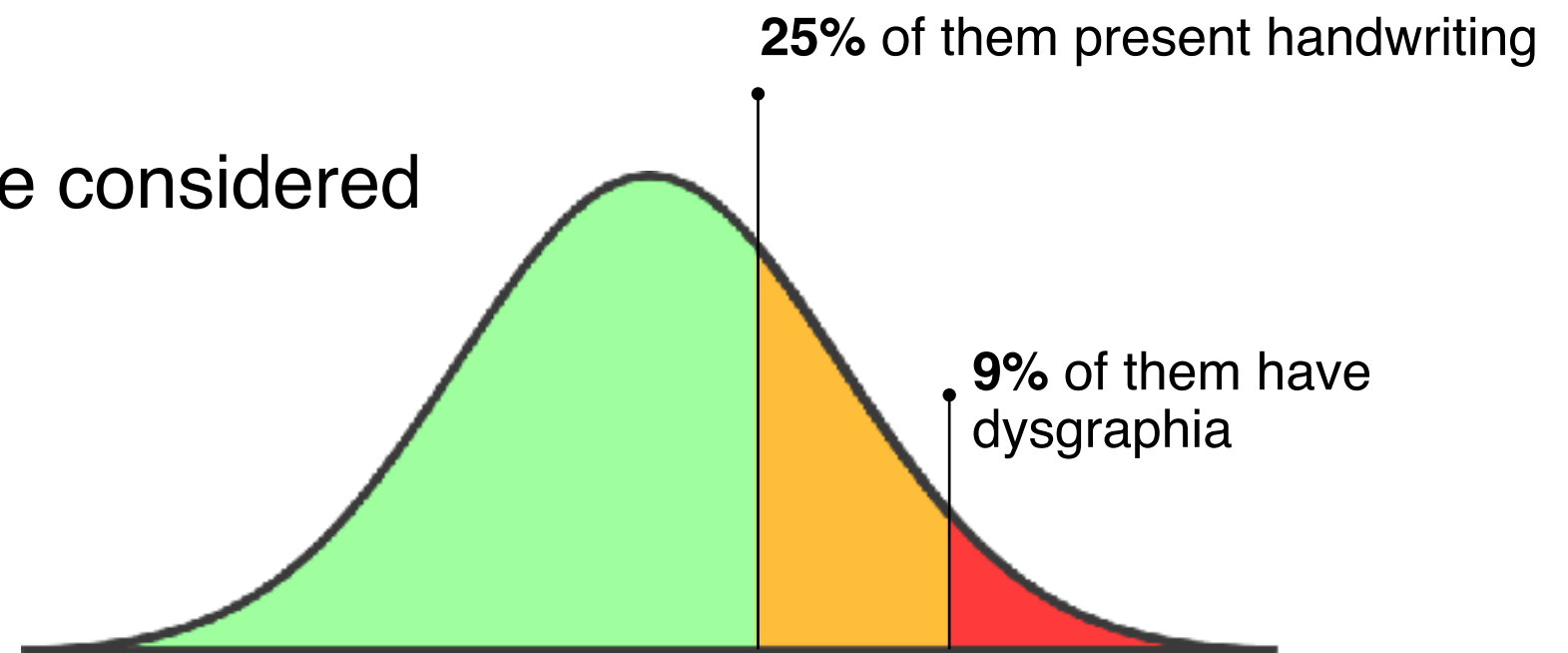
Domain-independent Intelligent Tutoring System

What are handwriting difficulties?

Handwriting difficulties are very common

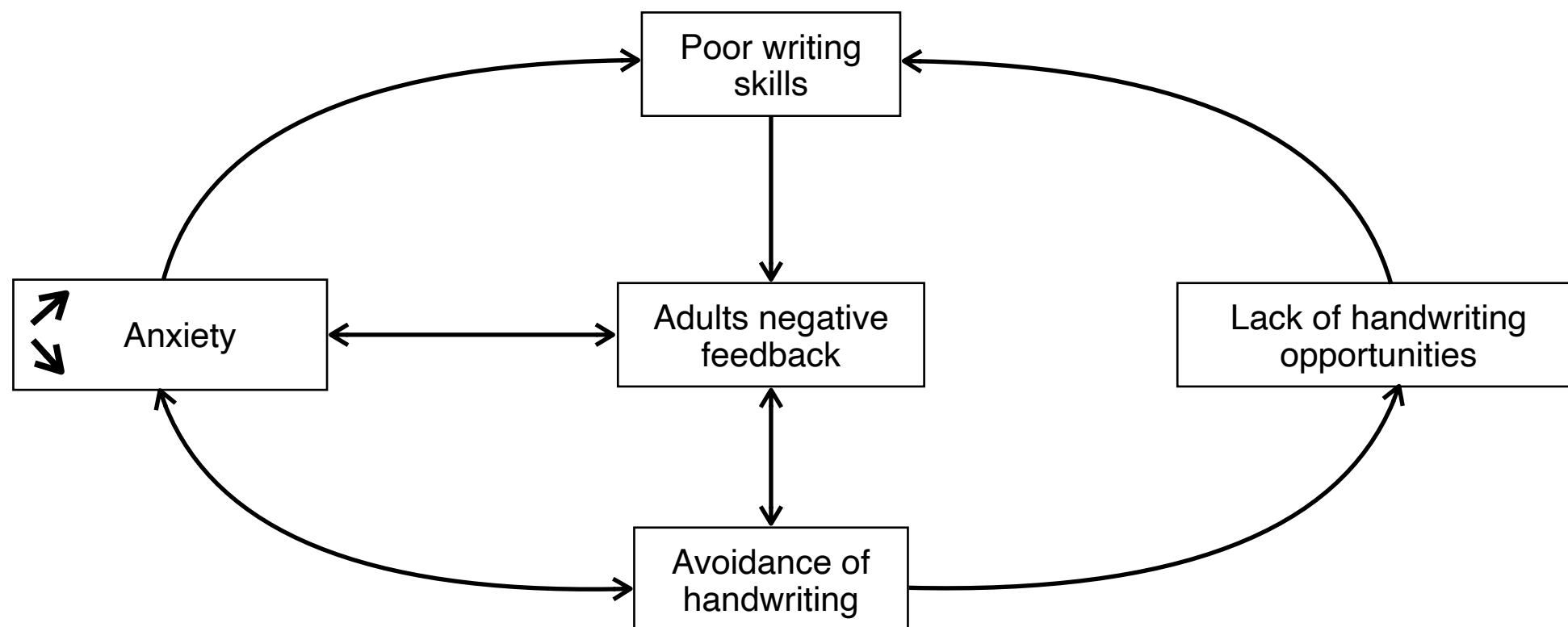
The definition and statistics about dysgraphia changes in function of the country.

In France, 8.6% of children are considered dysgraphic.



What are handwriting difficulties?

Critical consequences on the child's school curriculum but also on his/her behaviour



How handwriting difficulties are detected?

The BHK test

Il fait très beau
Je suis bien
Je vois de l'eau
Mais je ne sais pas
Où elle va.

CRITÈRES						TOTAL
1. Écriture grande						
2. Inclinaison de la marge vers la droite						
Phrases	1	2	3	4	5	
3. Lignes non planes						
4. Mots serrés						
5. Fermeture chaotique						
6. Liens interrompus entre les lettres						
7. Télescopages						
8. Variation dans la grandeur des lettres troncs						
9. Hauteur relative incorrecte						
10. Distorsion des lettres						
11. Formes de lettres ambiguës						
12. Lettres retouchées						
13. Hésitations et tremblements						
Score Total						

BHK: a test full of limitations

A test that needs to be adapted in the era of digitalization



Subjectivity

Relies on human judgement.

Small (60%) inter-rater agreement.



Time-consuming

Takes 5 minutes for children to pass the test.

Takes 15 minutes to correct it.



Costly

Experts need to be trained to acquire the expertise for grading.

Grading takes time.



Neglect information

Test done on paper:
All the dynamic, pressure and tilt characteristics of handwriting are not perceivable.

The dynamics of handwriting is important

The database

- More >1000 children asked to write cursively the 26 letters of the alphabets + the 10 digits.
- Data acquired with a Wacom tablet allowing to record the dynamic of the movement.

School children

971 children from 14 schools, from pre-school (*5 years old*) to 5th grade (*10 years old*)

Dysgraphic children

27 children (diagnosed dysgraphic according to the BHK test)

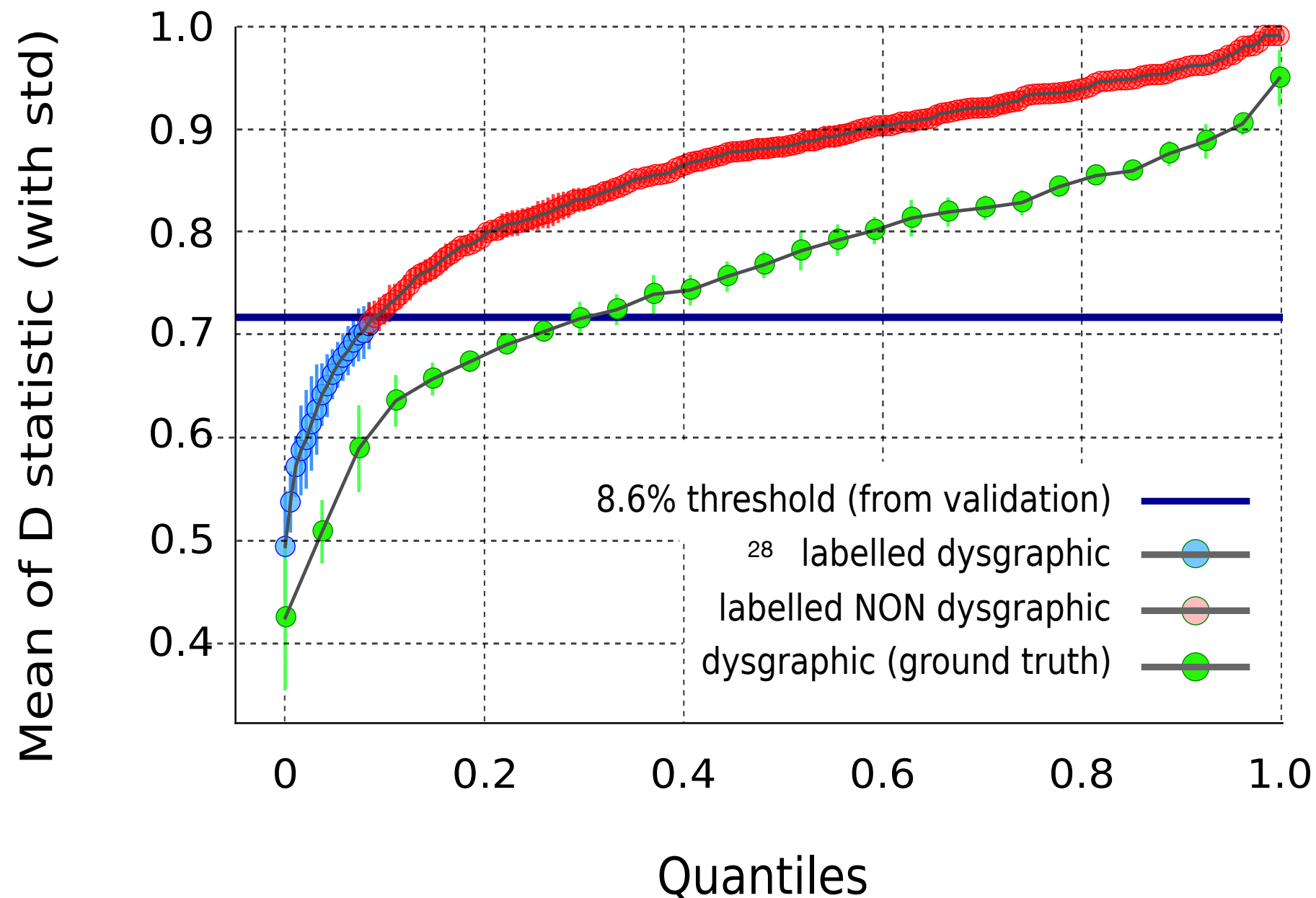


The dynamics of handwriting is important

Training a CNN to detect handwriting difficulties

The Convolutional Neural Network (CNN) takes the final output (image) of handwriting as an input...

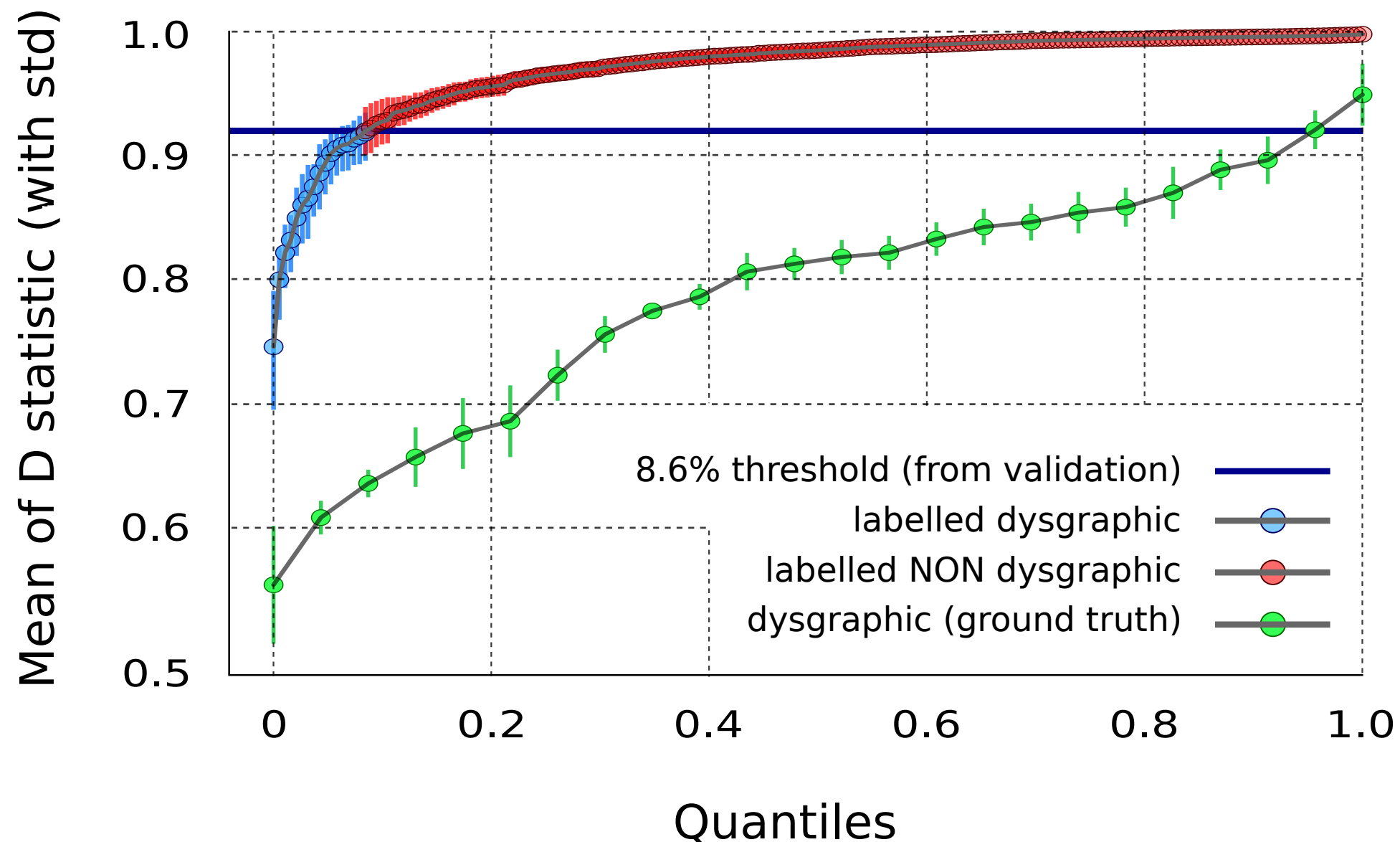
... and fails to accurately discriminate dysgraphic and non dysgraphic children.



The dynamics of handwriting is important

Training a RNN to detect handwriting difficulties

The Recurrent Neural Network (RNN) captures the dynamic of handwriting...
... and accurately discriminates dysgraphic and non dysgraphic children.



A supervised learning method to detect dysgraphia

The database

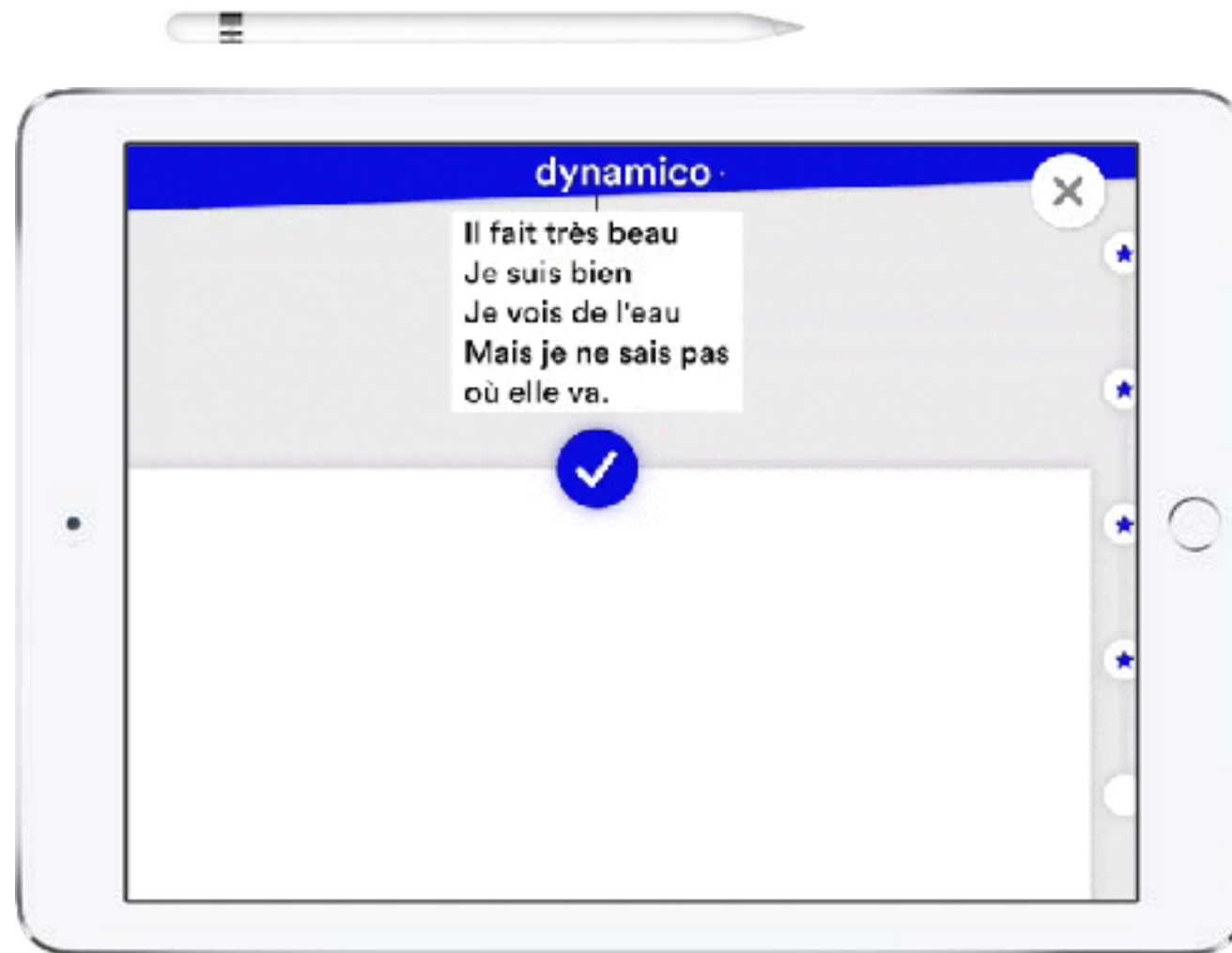
- 298 children asked to write the BHK test.
- Data acquired with a Wacom tablet allowing to record the dynamic of the movement, but also the pressure and the tilt of the pen.

School children

242 children from 14 schools, from pre-school (*5 years old*) to 5th grade (*10 years old*)

Dysgraphic children

56 children (diagnosed dysgraphic according to the BHK test)



Static



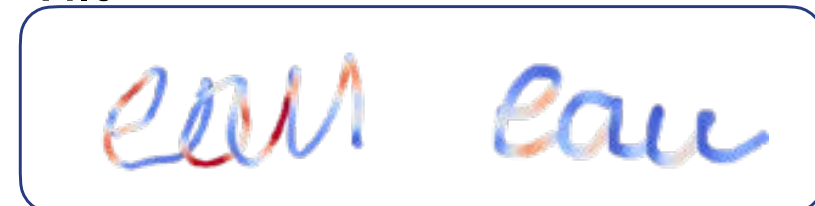
Kinematic

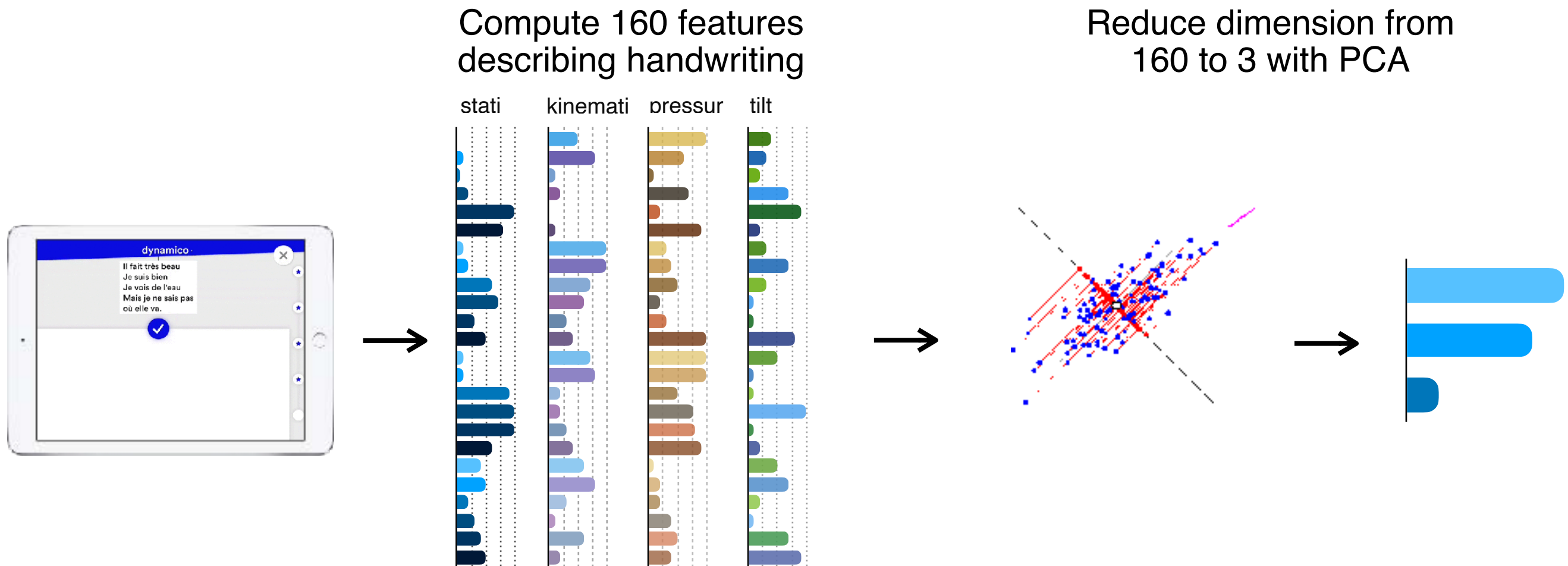


Pressure

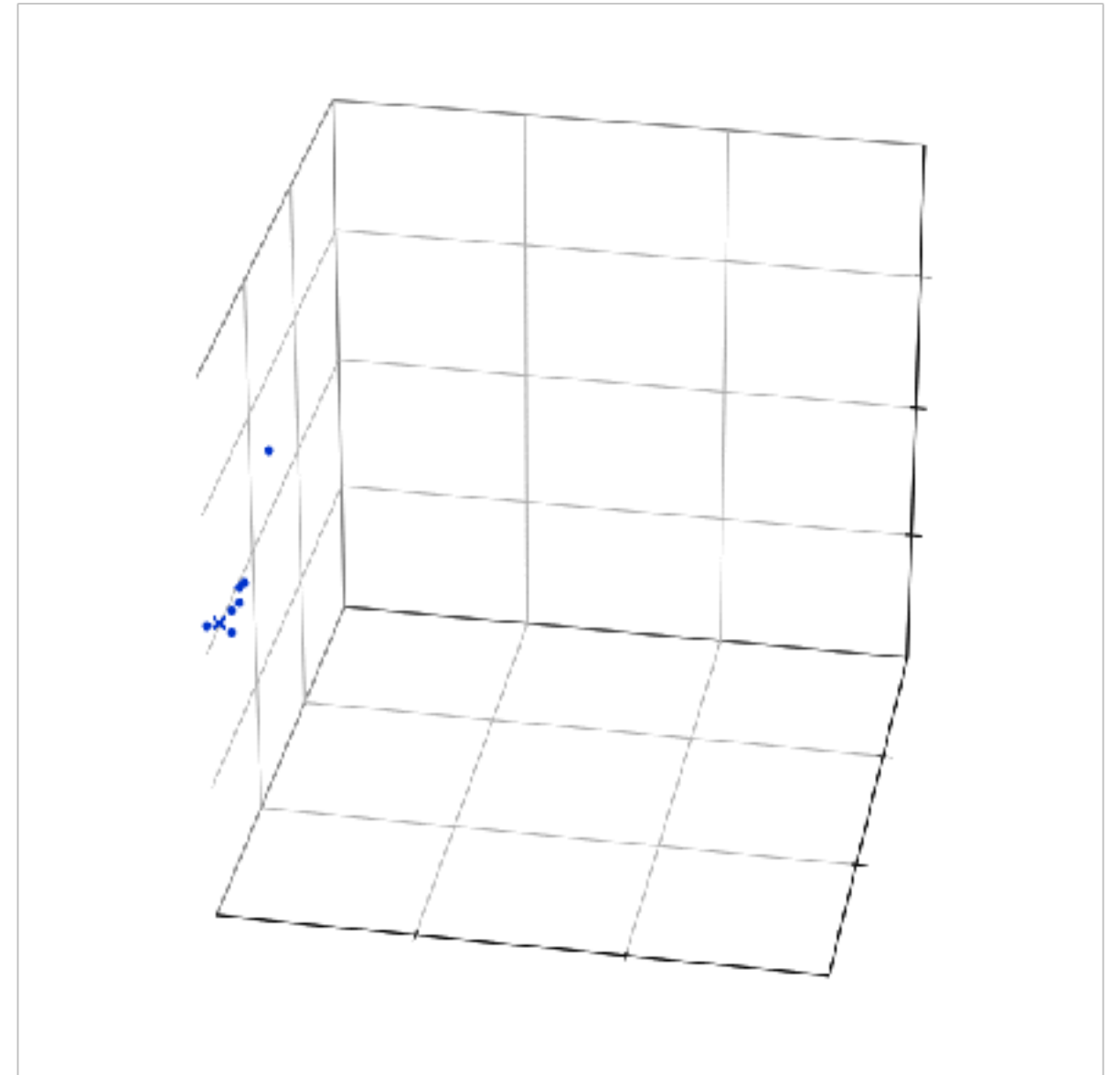


Tilt



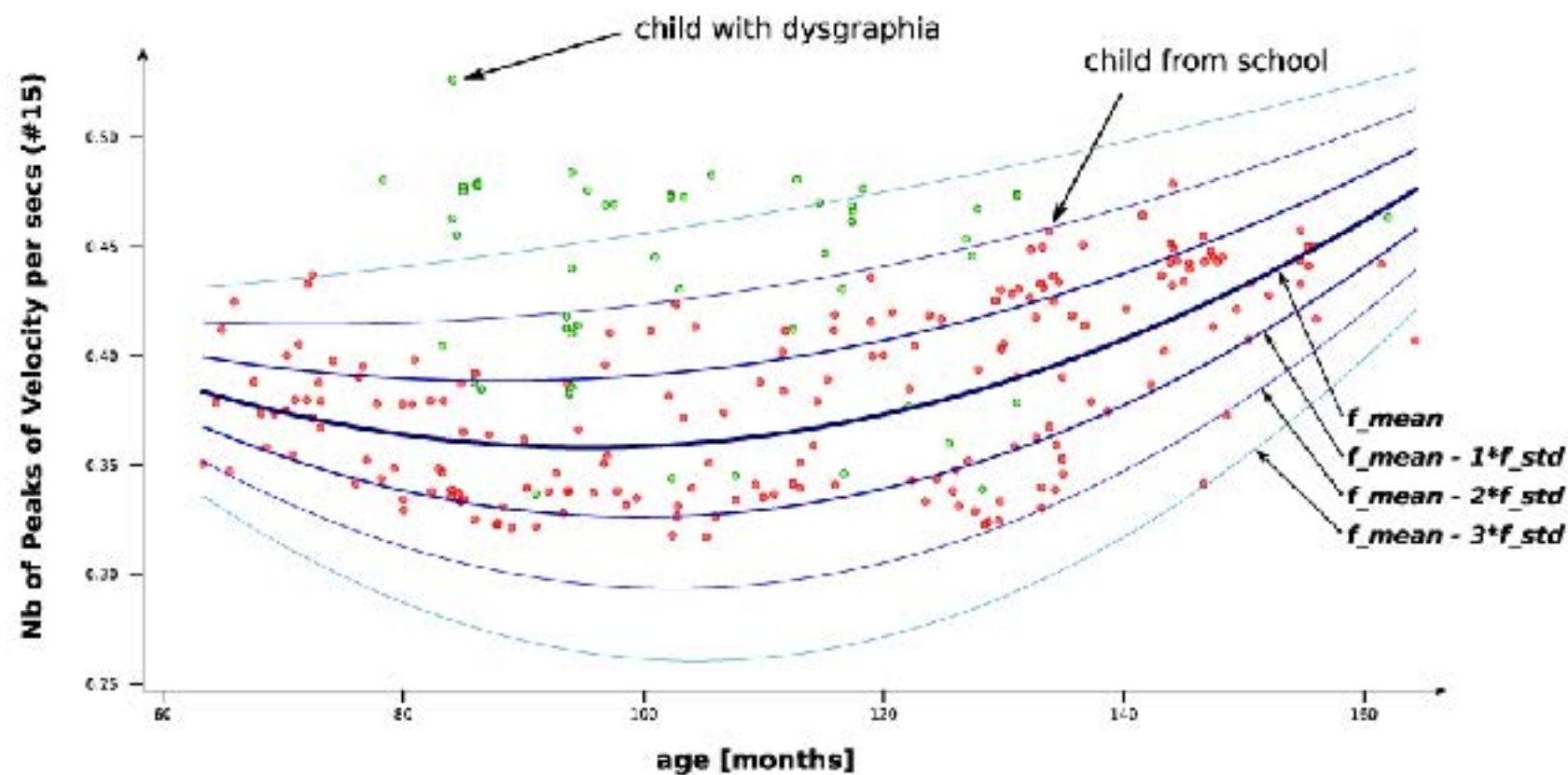


- Unsupervised machine learning to find two clusters (dysgraphic and non dysgraphic children).
- More than 90% of dysgraphics clustered together.
- 12% of children from school detected as dysgraphic, in line with the statistics (8.6%).



An unsupervised method to analyse handwriting

- Handwriting evolves with age, we need to take care of this.
- Different models for each gender, and each country.



The most important features used by the Random Forest to detect dysgraphia sorted.

Table 2. The most important features found by the Random Forest model, using Gini importance as a metric

Rank	Category	Name	Importance (Std.) [%]
1	Kinematic	Median of Power Spectral of Speed Frequencies	15.71 (9.06)
2	Kinematic	Bandwidth of Speed Frequencies	12.08 (8.00)
3	Pressure	Mean Speed of Pressure Change	9.81 (6.52)
4	Static	Space Between Words	7.45 (6.73)
5	Tilt	Distance to Mean of Speed of Tilt-X Change Frequencies	6.07 (4.30)
6	Kinematic	Distance to Mean of Speed Change Frequencies	5.18 (4.73)
7	Tilt	Bandwidth of Speed of Tilt-X Change Frequencies	4.10 (4.64)
8	Tilt	Median of Power Spectral of Tilt-Y Change Frequencies	2.97 (3.33)

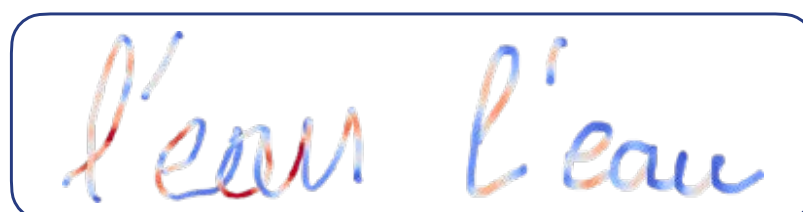
We report the ranks, feature categories, and their importance averaged for the 25 folds and the standard deviation of importance over all folds

Only one of these features could be extracted with a pen/paper test!

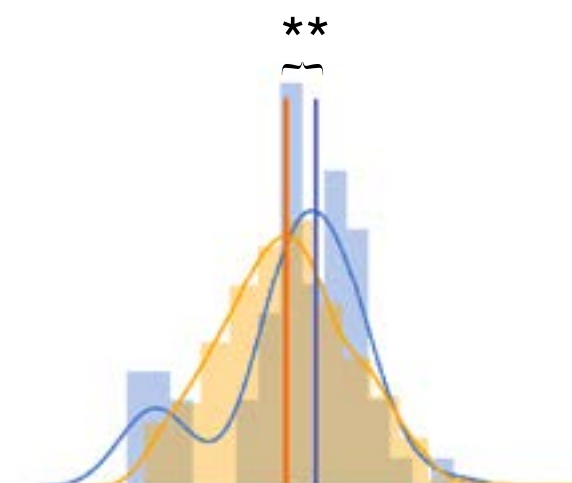
Static



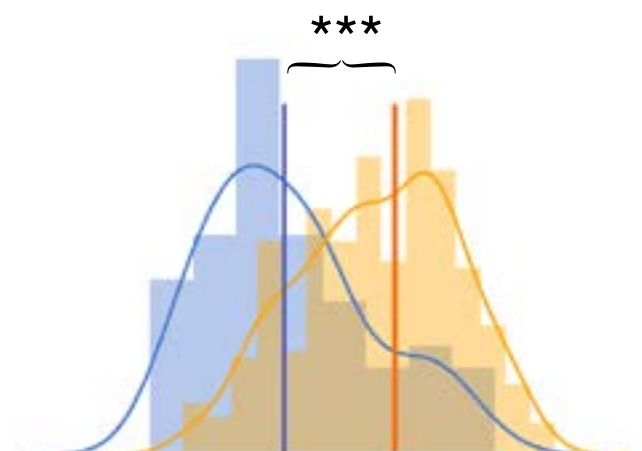
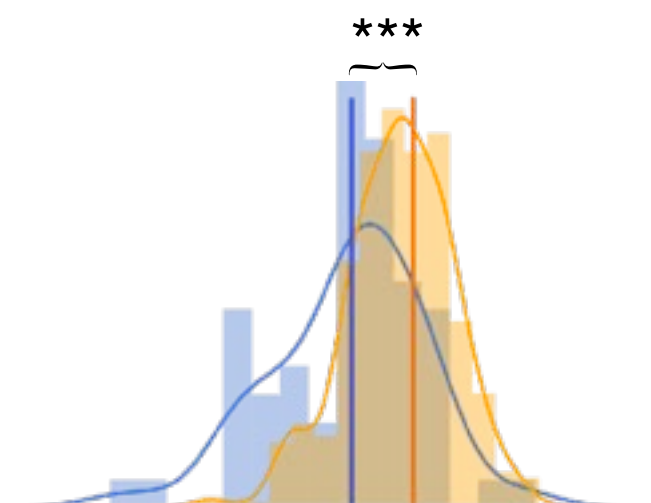
Kinematic



Pressure

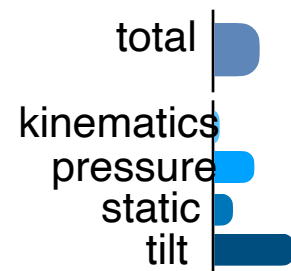
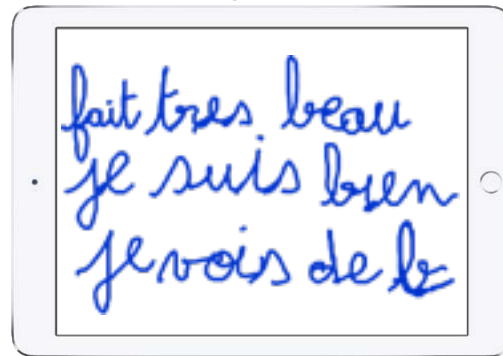


moment of handwriting

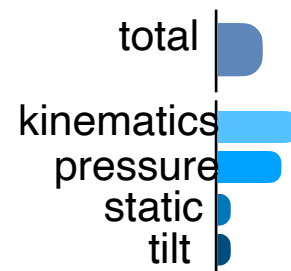
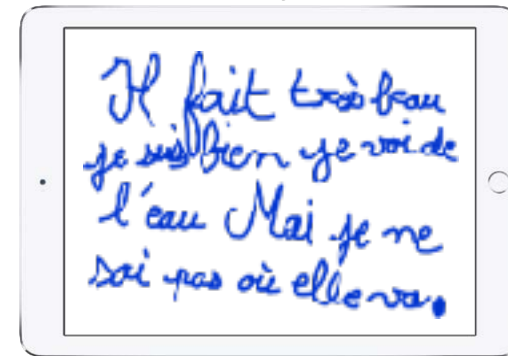
median of power spectral
of speed frequenciesmedian of power spectral
of speed of pressure
change frequencies

The difficulties are not the same for all kids

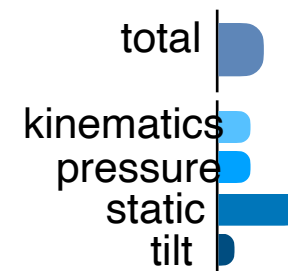
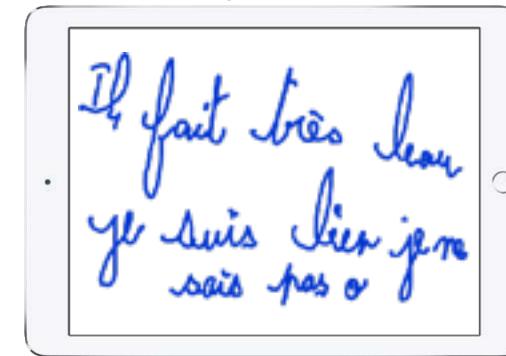
Marc, 8 y.o



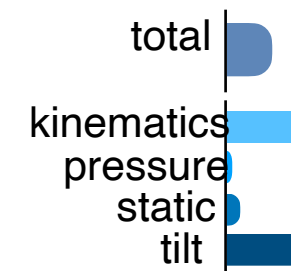
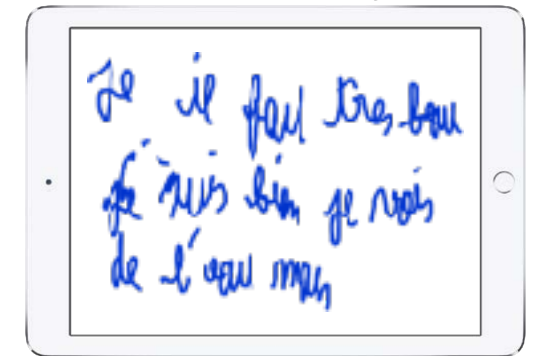
Patrick, 5 y.o



Julia, 6 y.o



Stephanie, 7 y.o

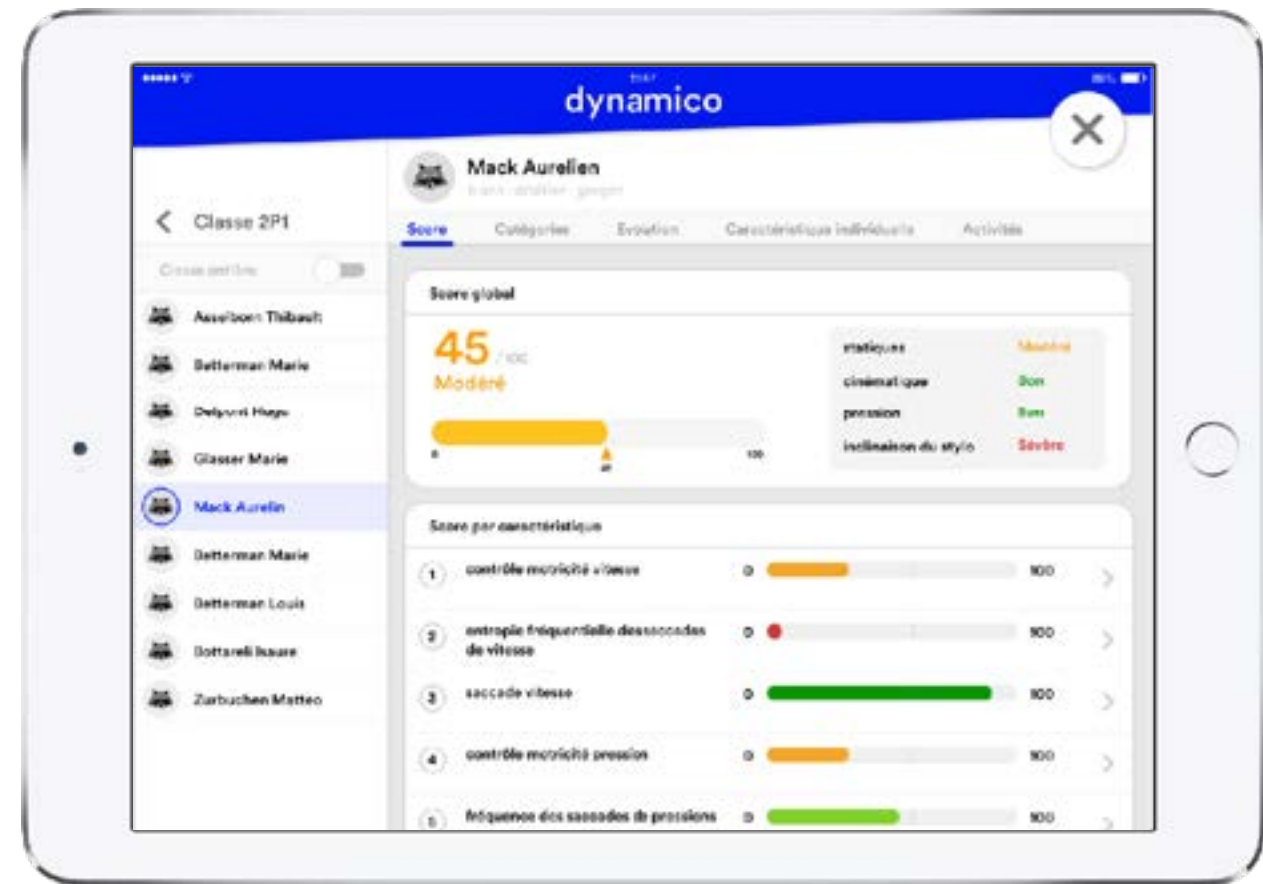
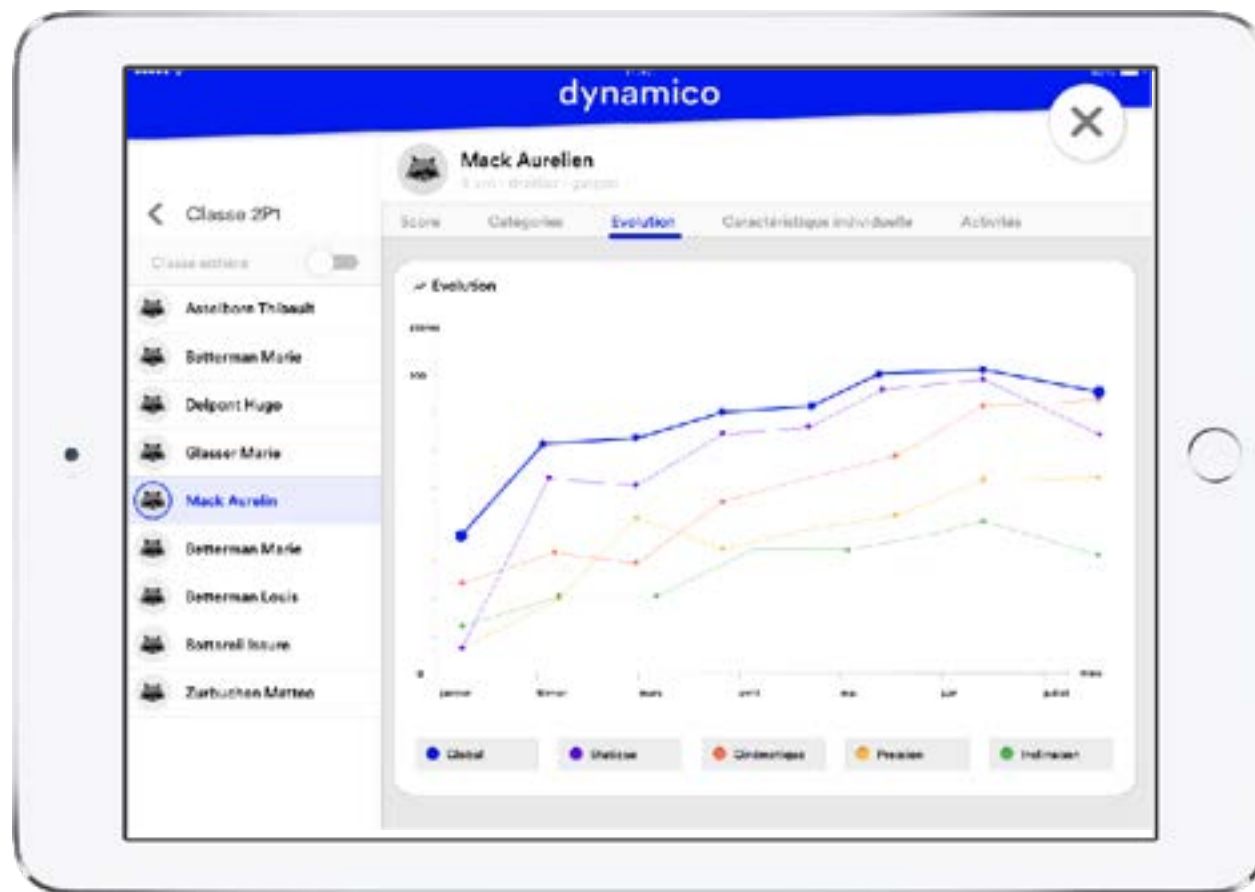


Jeux spécifiques pour s'améliorer



dynamilis

Suivi des élèves



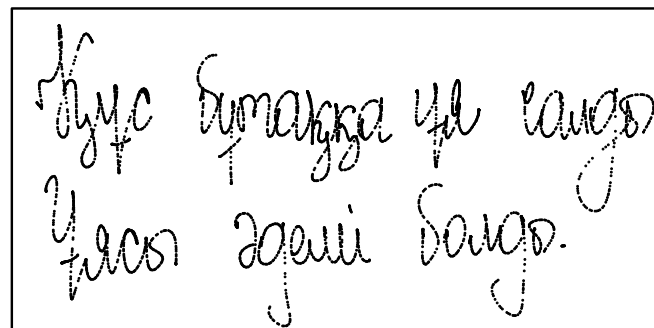
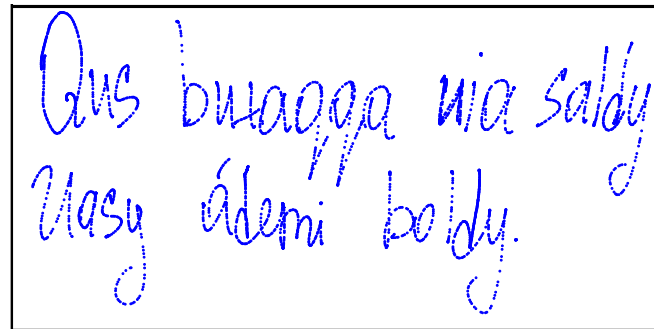
Transferring the handwriting analysis model

Can we transfer the model to other alphabets or supports?

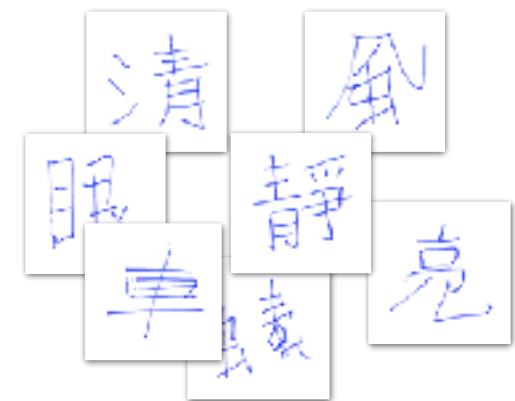
Paper vs tablet



Latin vs Cyrillic alphabet



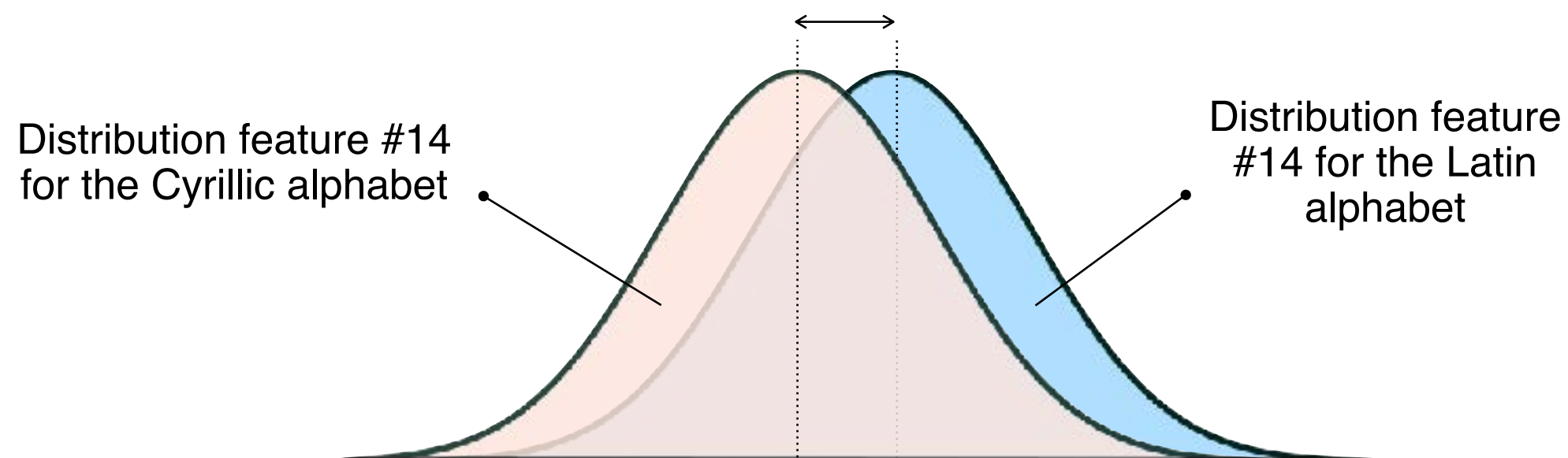
English vs Chinese alphabet



Asselborn et al., The transferability of handwriting skills: from the Cyrillic to the Latin alphabet, 2020, *npj Science of Learning* (under review)

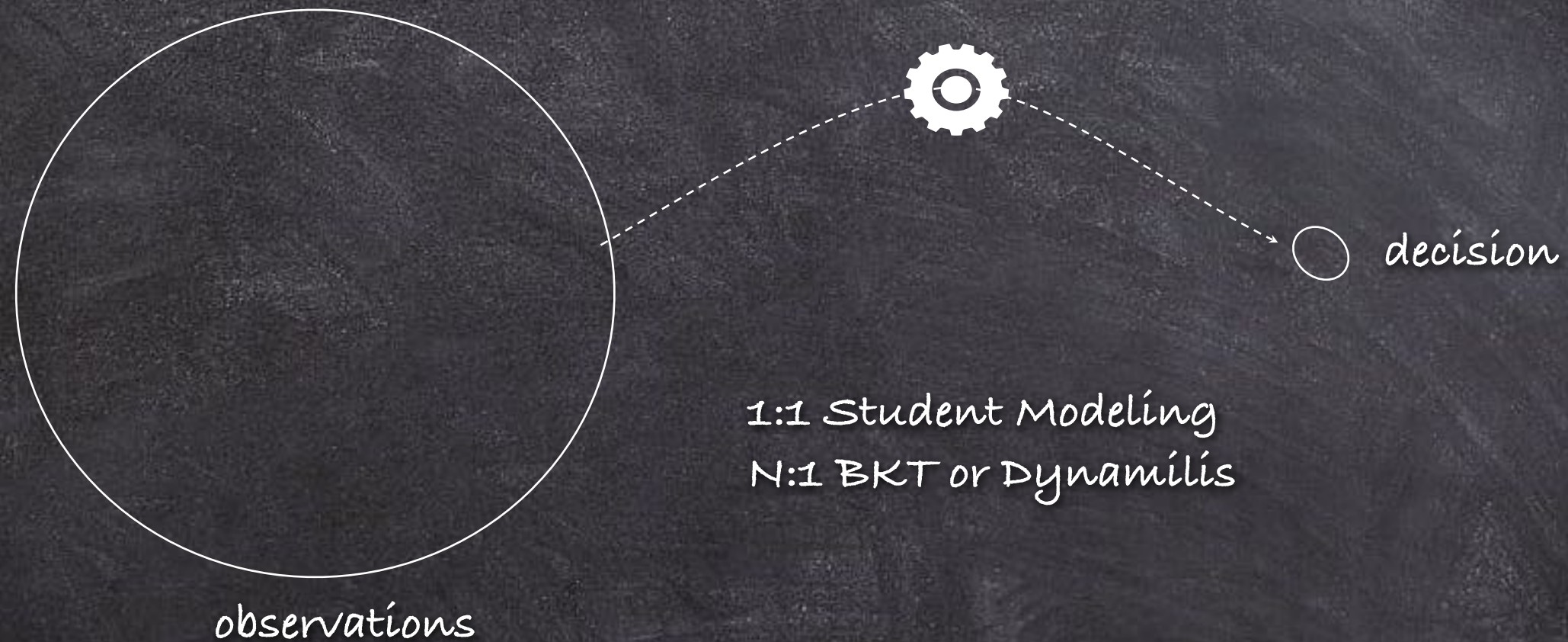
Transferring the handwriting analysis model

Is there an absolute difference between the features' distribution?

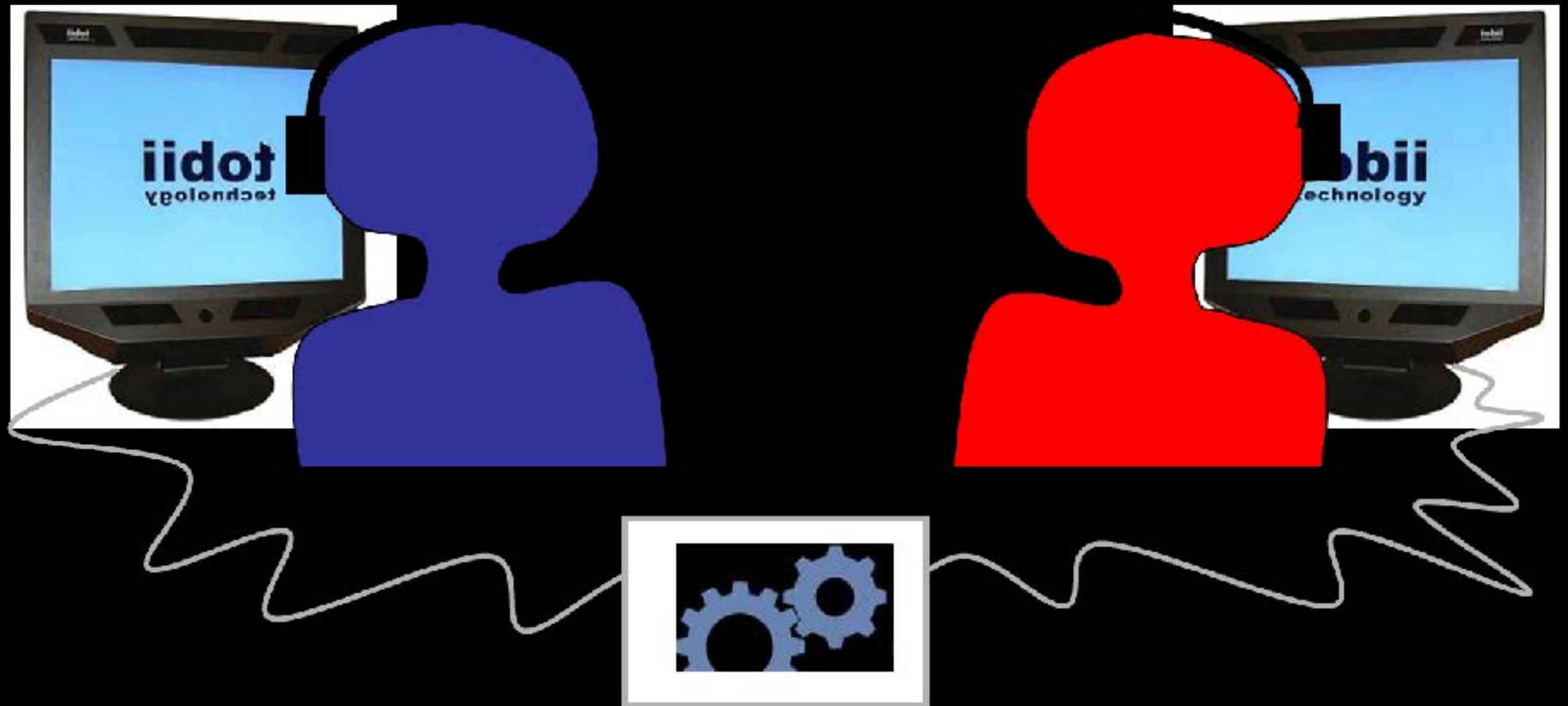


~~General~~ AI

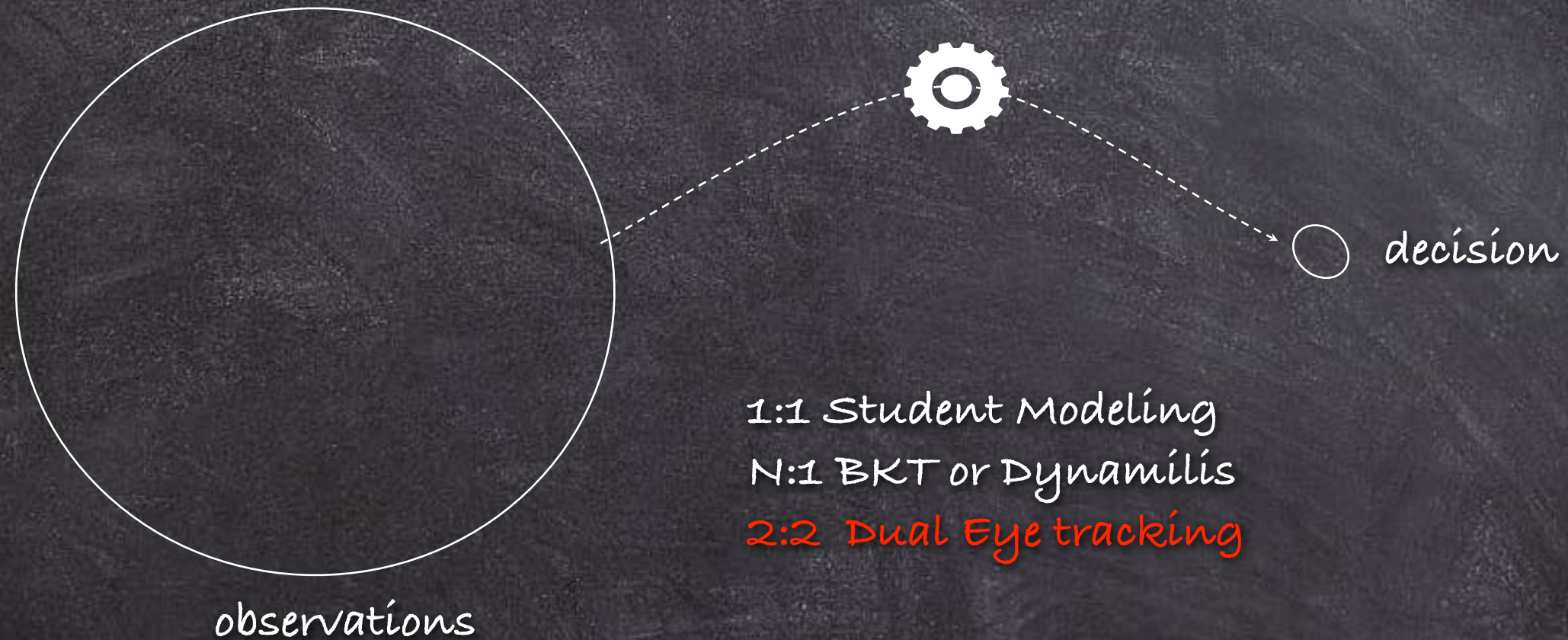
Learning Analytics Ratios



Dual Eye Tracking



Learning Analytics Ratios



DUET - Dual Eye-Tracking
Pair programming experiment

Low gaze recurrence



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

P. Jermann, M. A. Nüssli & P. Dillenbourg
© CRAFT - <http://craft.epfl.ch/>

Supported by the Swiss National Science Foundation
(grants #K-12K1-117909 and #PZ00P_126611)

DUET - Dual Eye-Tracking
Pair programming experiment

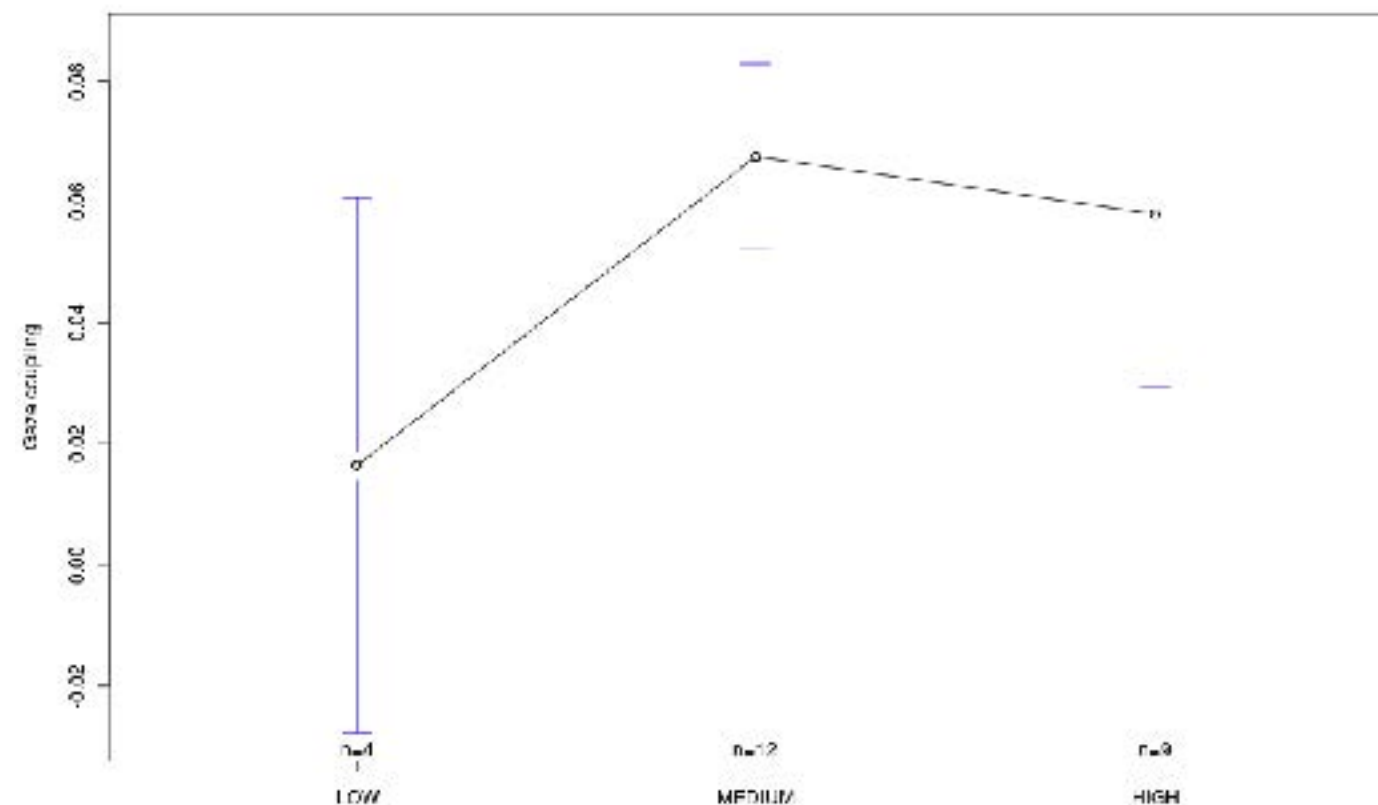
High gaze recurrence

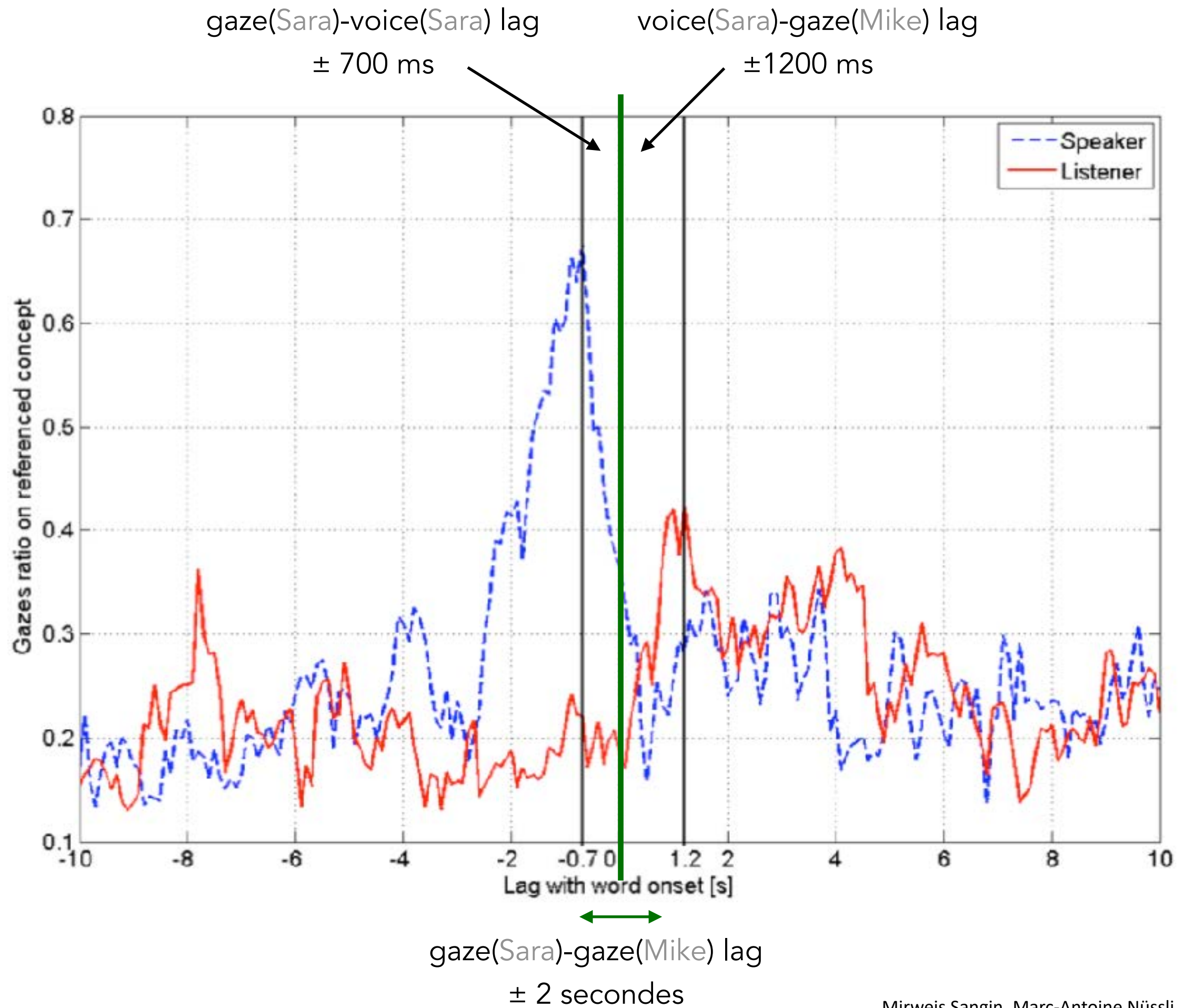


ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

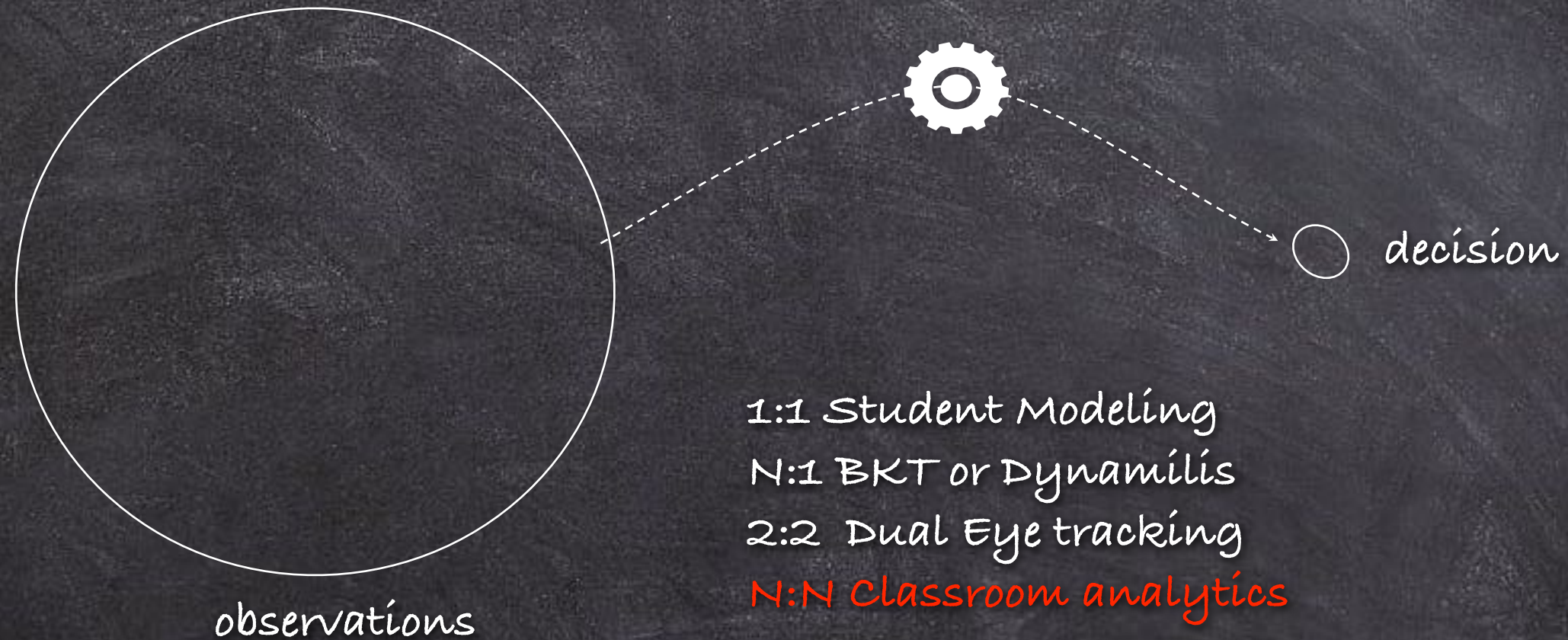
P. Jermann, M.-A. Nüssli & P. Dillenbourg
© CRAFT - <http://craft.epfl.ch/>

Supported by the Swiss National Science Foundation
(grants #K-12K1-117909 and #PZ00P_126611)





Learning Analytics Ratios

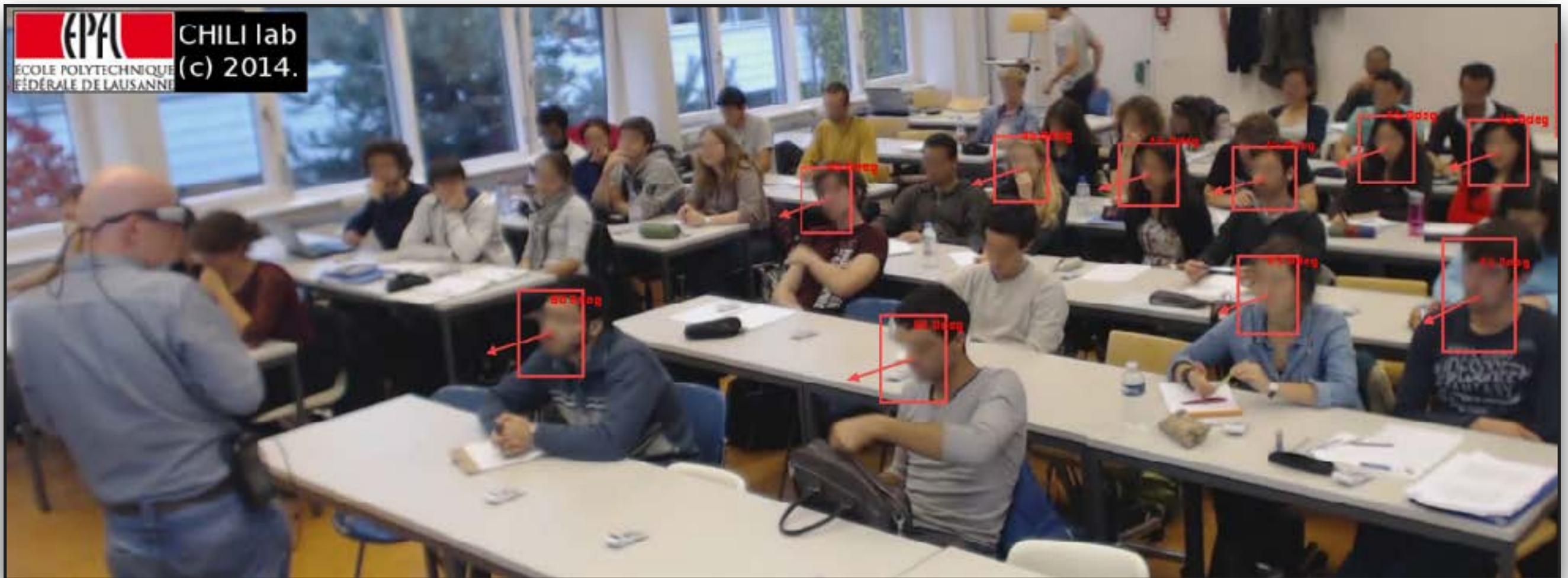




You passed below 5% attention



Classroom Analytics



Do you really want to put cameras in classes ?

NO

EPFL Exercises Session

assistant

works



waits



“While Waiting Productivity” LOSS : 62% → 6%



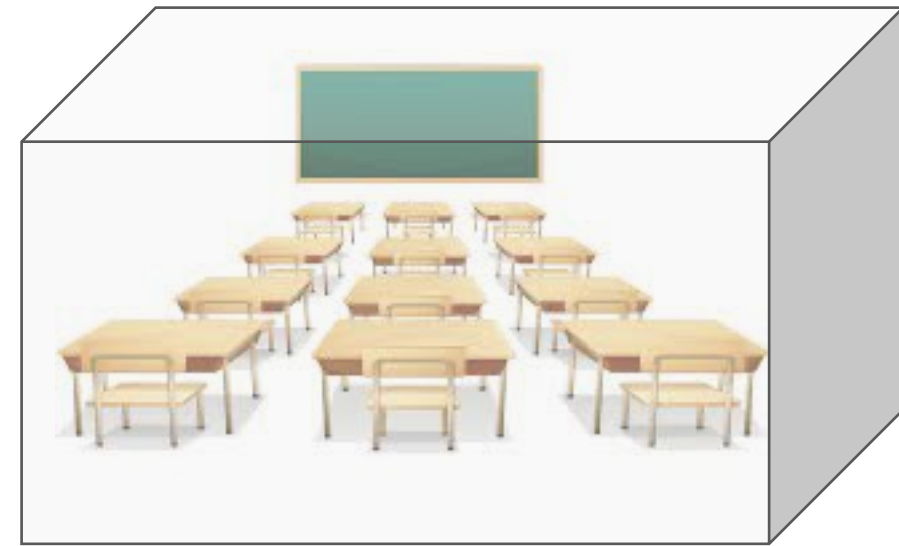
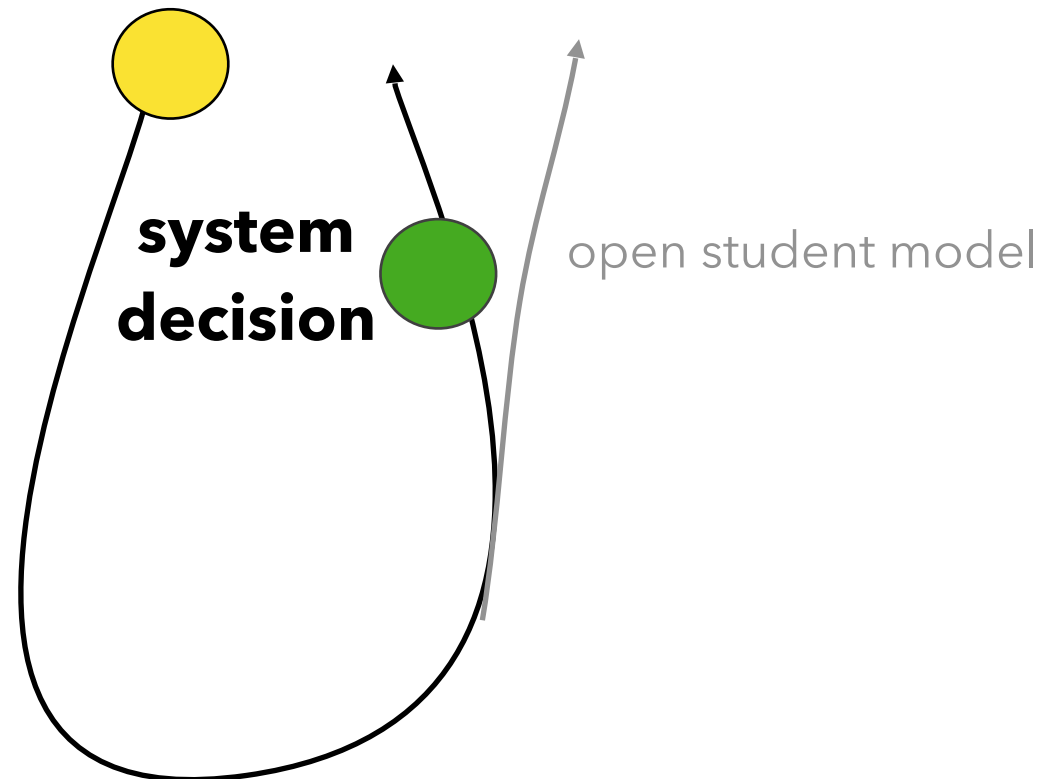
Classware: the classroom is a digital system





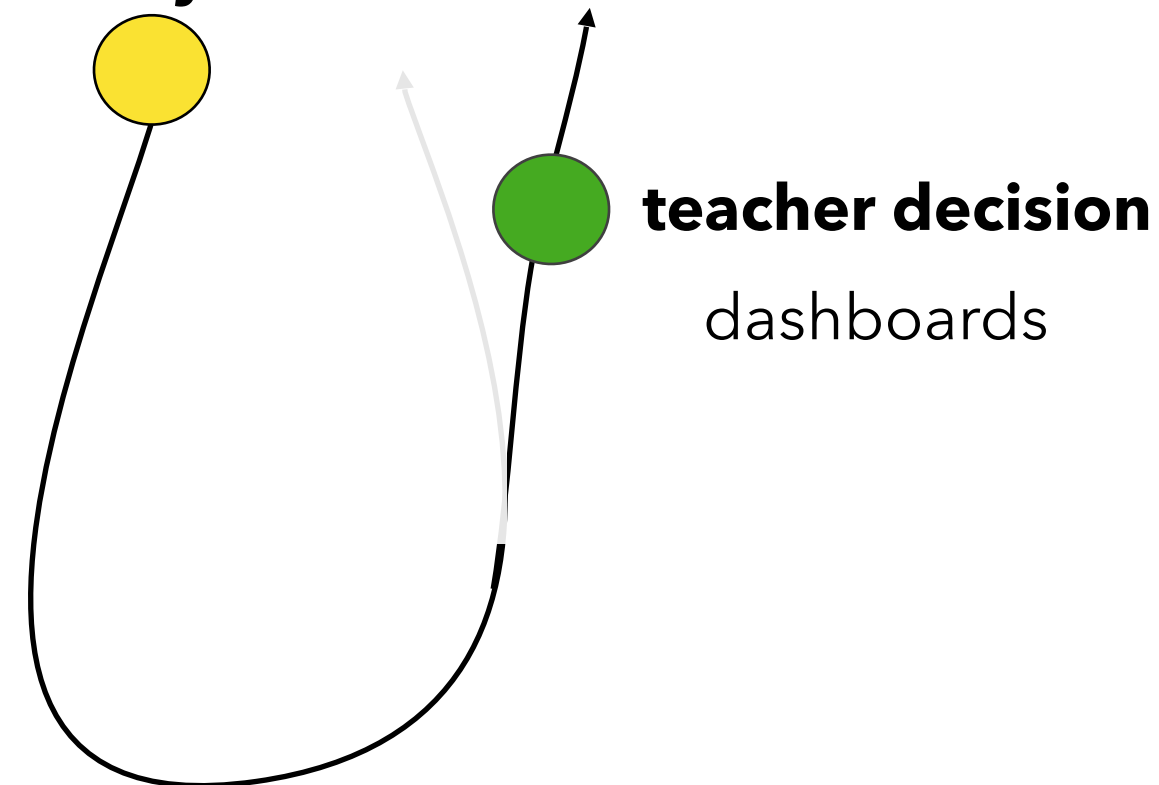
learning analytics

individual student
knowledge state

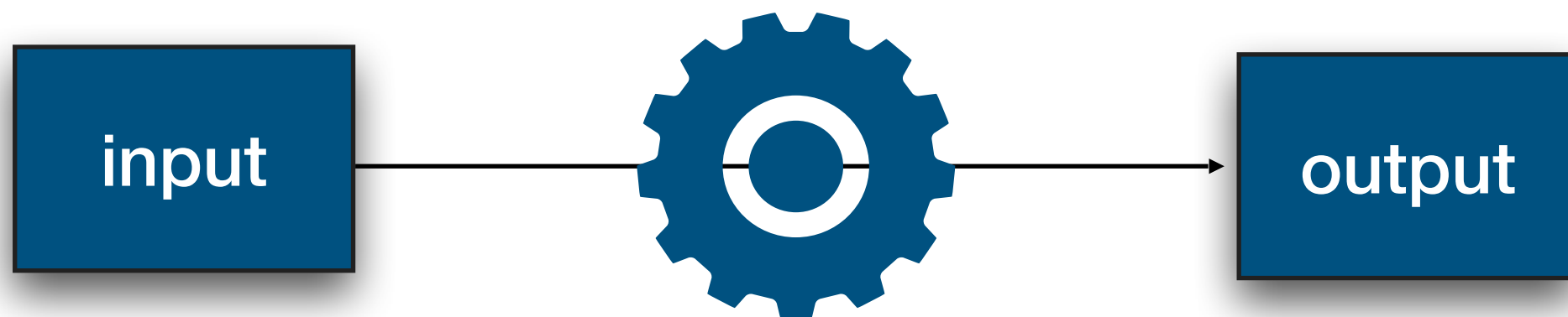


classroom analytics

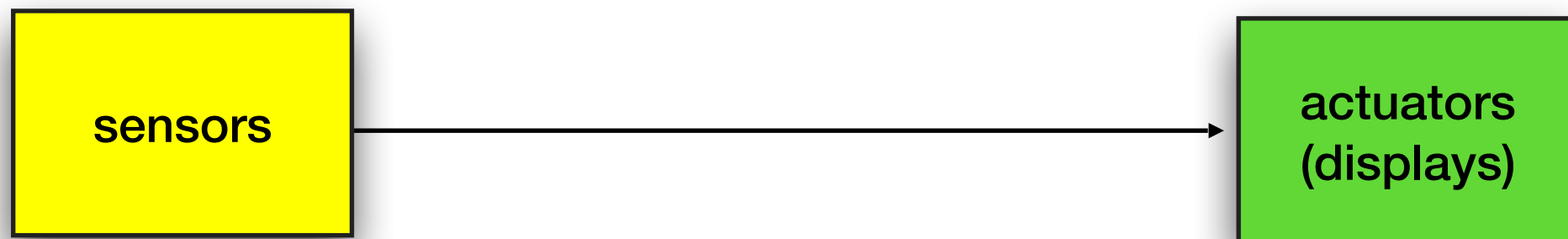
n students
activity state



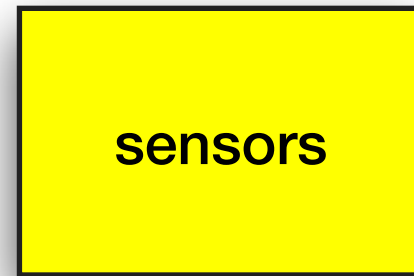
My (physical) classroom is a *digital* environment

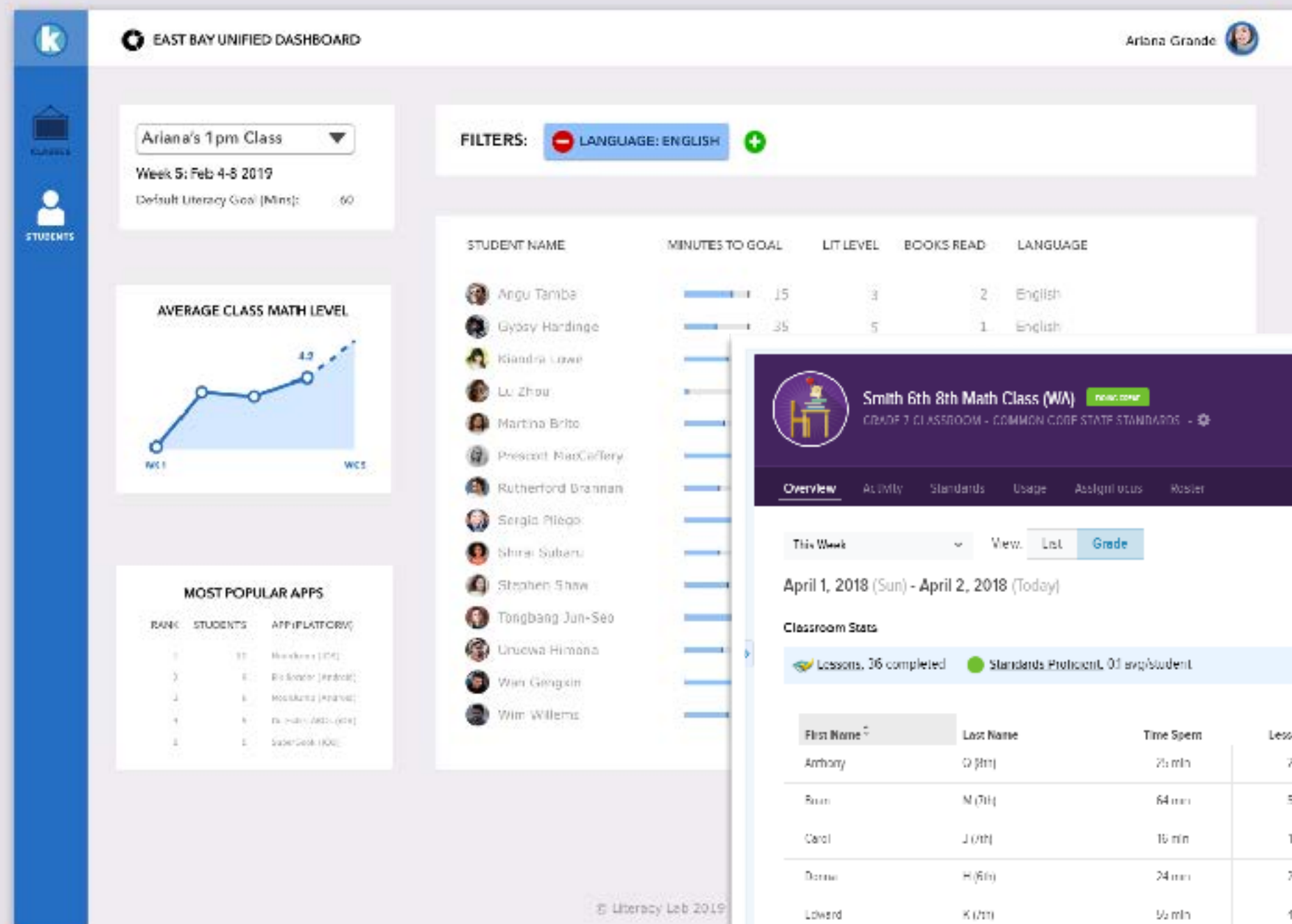


My (physical) classroom is a digital environment



- **Accelerometer** that measures acceleration forces
 - **Magnetometer** that measures magnetization/magnetic fields
 - **Light sensor** that measures the strength of light
 - **Gyroscope** that measures orientation
 - **Hygrometer** that measures moisture in the atmosphere
 - **Thermometer** that measures ambient temperature
 - **Barometer** that measures atmospheric pressure
 - **Proximeter** that detects when something is close
-
- **Precision Gas Sensor** - Test air quality, carbon monoxide levels
 - **Reducing Gas Sensor** - Sense methane, propane, and natural gas leaks
 - **Oxidizing Gas Sensor** - Ozone sensing and chlorine leaks
 - **Non-Contact Thermometer** - Check food temperature, engine diagnostics, ...
 - **Humidity Sensor** - Check heat index, food storage conditions
 - **Temperature Sensor** - Measure ambient temperature
 - **Light Sensor** - For checking light intensity, sunlight monitoring,....
 - **Color Sensors** - Use as a color meter, color matcher/analyzer, even pattern recognition
 - **Pressure Sensor** - Blood pressure monitor
 - **Proximity Sensor** - Use it as a stud finder or liquid level monitor
 - **Expansion Connector** - Connecting other sensors like EKGs, Thermal Printers, and more



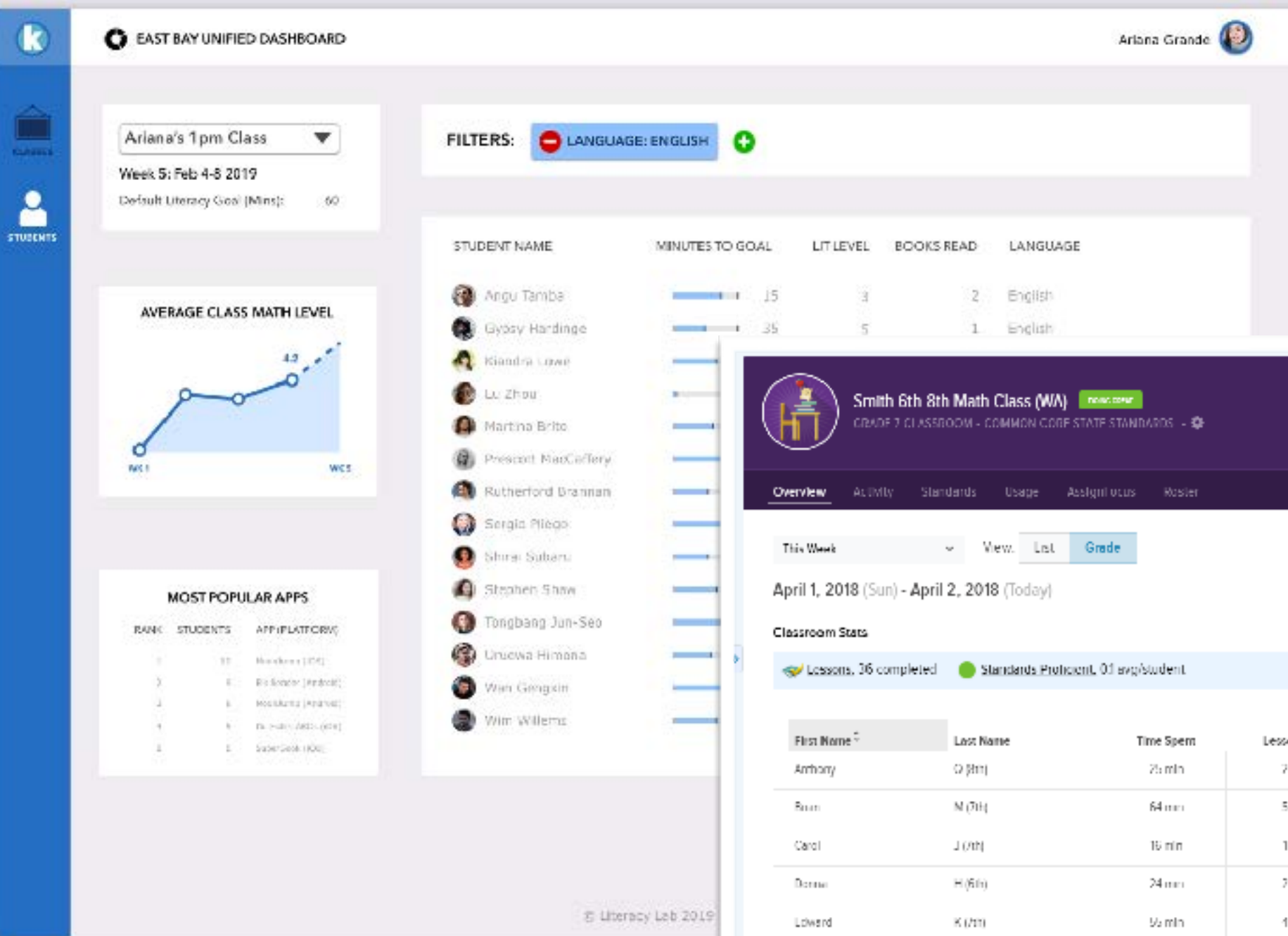


Kidapolis

Dreambox

Dashboard

actuators
(displays)



Kidapolis

Dreambox

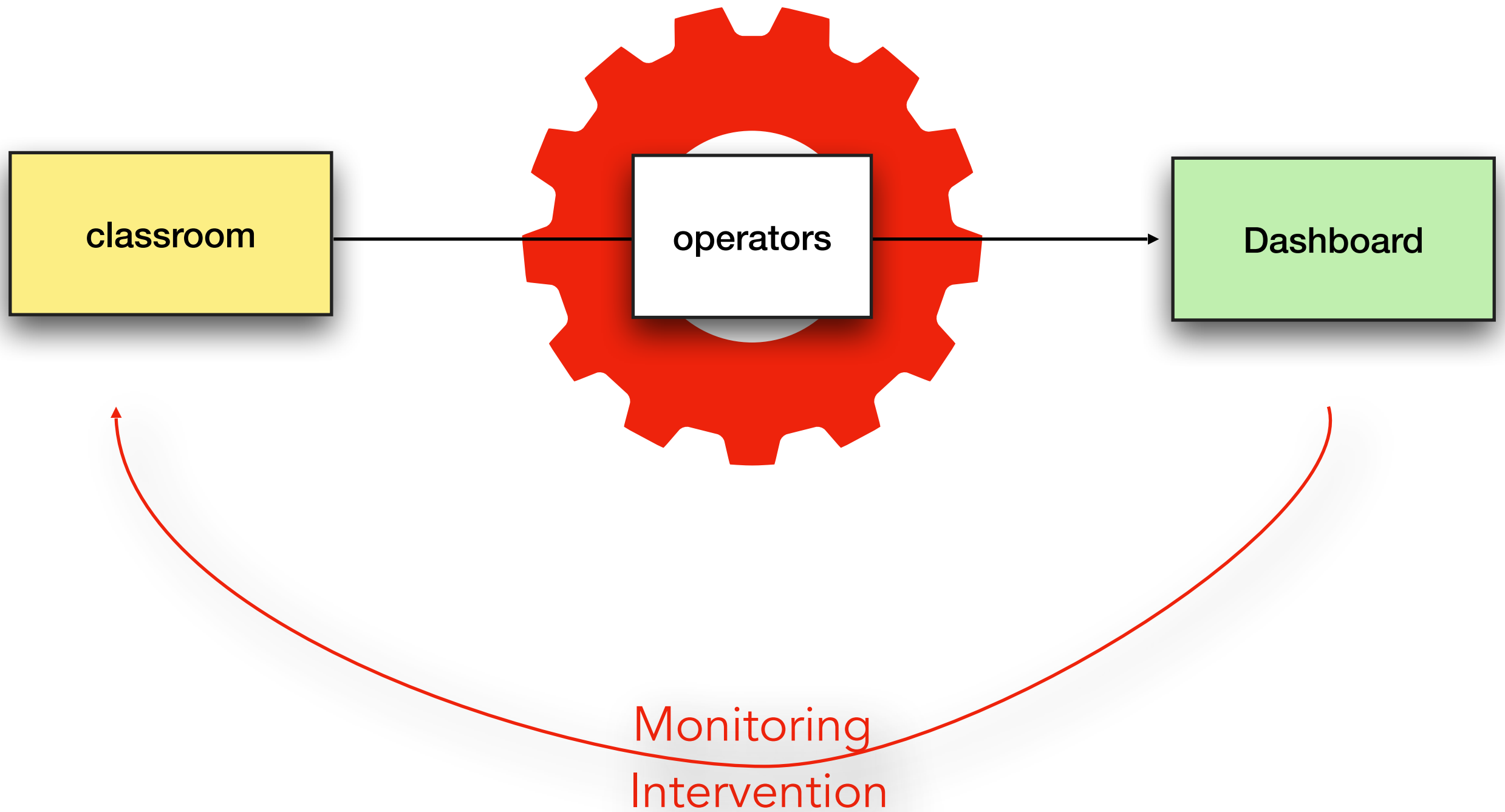
Dashboard

Classroom
Data

Visualisation



Aggregation, comparison, visualisation,....



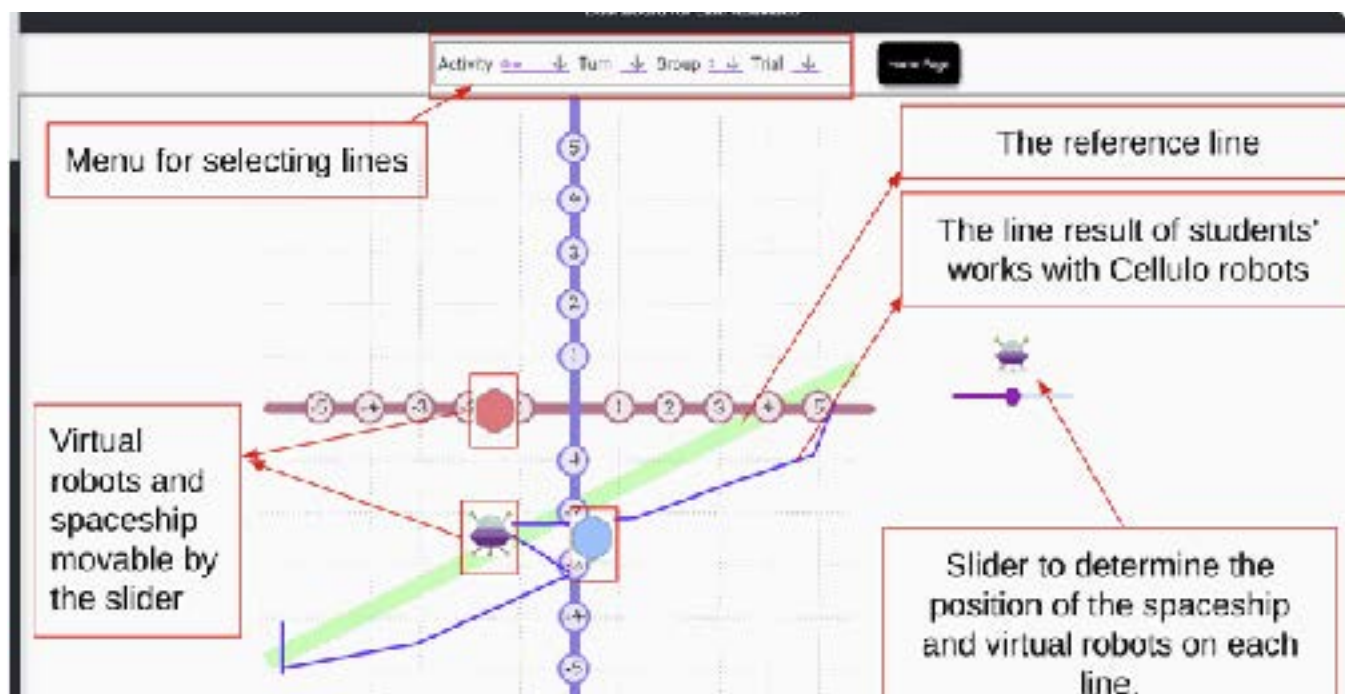


Sina Shahmoradi, EPFL

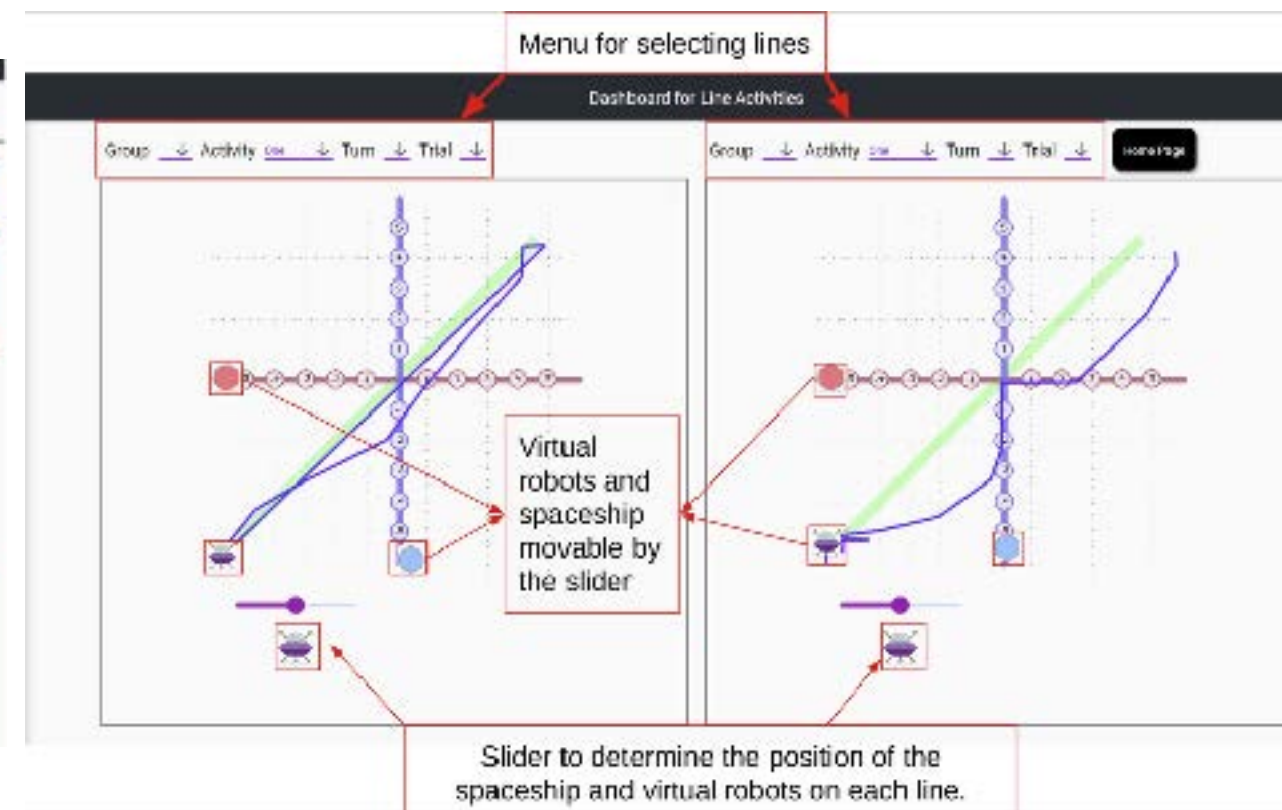
Teacher's Dashboard



Aggregated mode



Zoom mode



Comparison mode

Pause/Resume
the class activity

Progress-bar

Robot Failure Alert

Pause/Resume a
group activity

The robot of group 5 does not work.

	1) Warm up	2) Co-Captains	3) Alien Radar	4) Lost astronauts	
1 sina maria pierre	✓ ✓	✓ ✓ ✓	✓ ✓ ✓	■	⏸ ↗
2 dina yina net	✓ ✓ ✓	✓	✓ ✓ ✓	✓	⏸ ↗
3 kevin yannia guillaume	✓ ✓	✓ ■	✓ ✓ ✓	✓ ■	⏸ ↗
4 hala florence victor	✓ ✓ ■	✓ ✓ ✓	✓ ✓ ✓	✓ ■	⏸ ↗
5 jcun maryem foad	✓ ✓ ✓	✓ ✓ ✓	✓ ✓ ✓	✓ ✓	⏸ ↗



Distributed



Data



Visualisation

Centralized

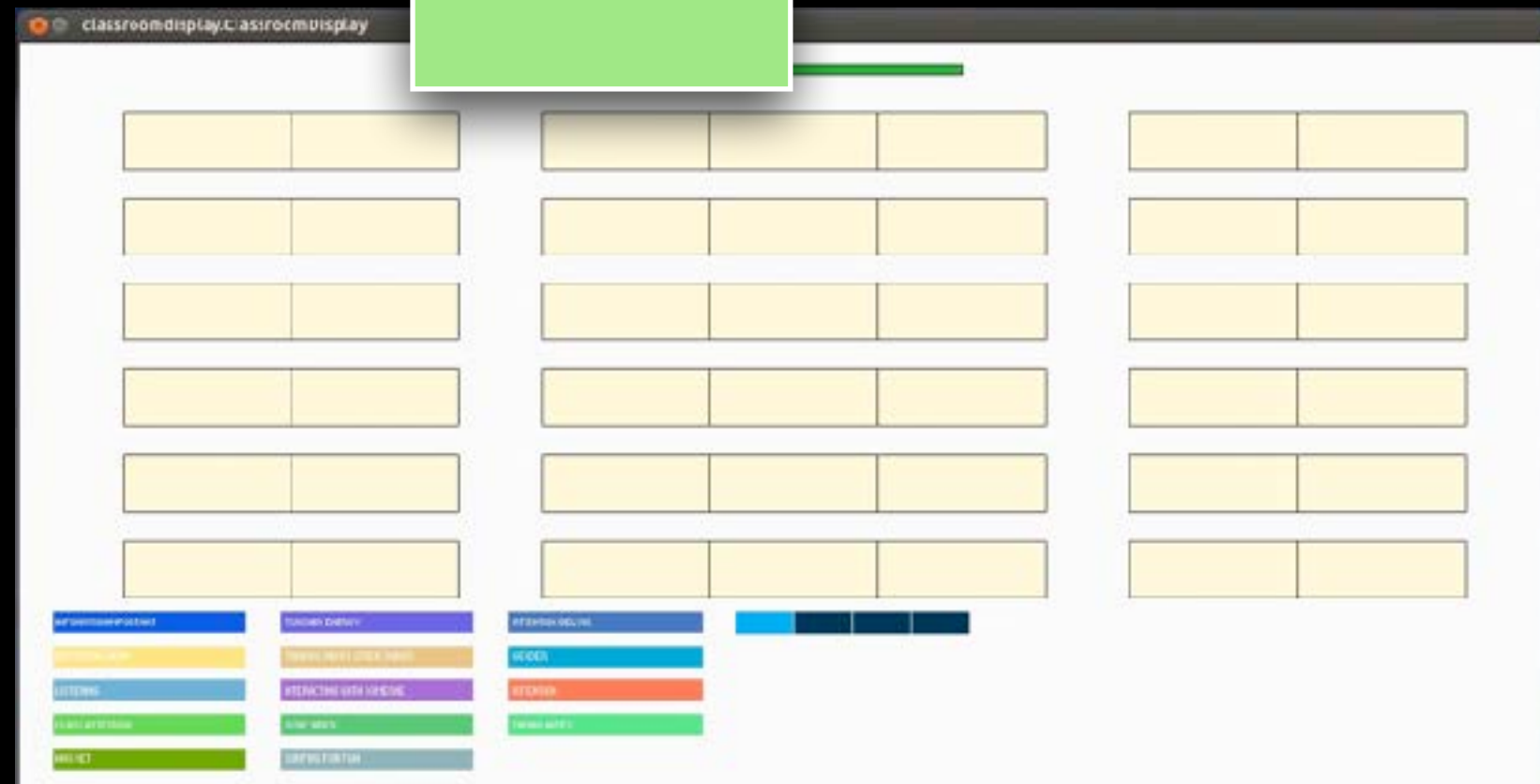
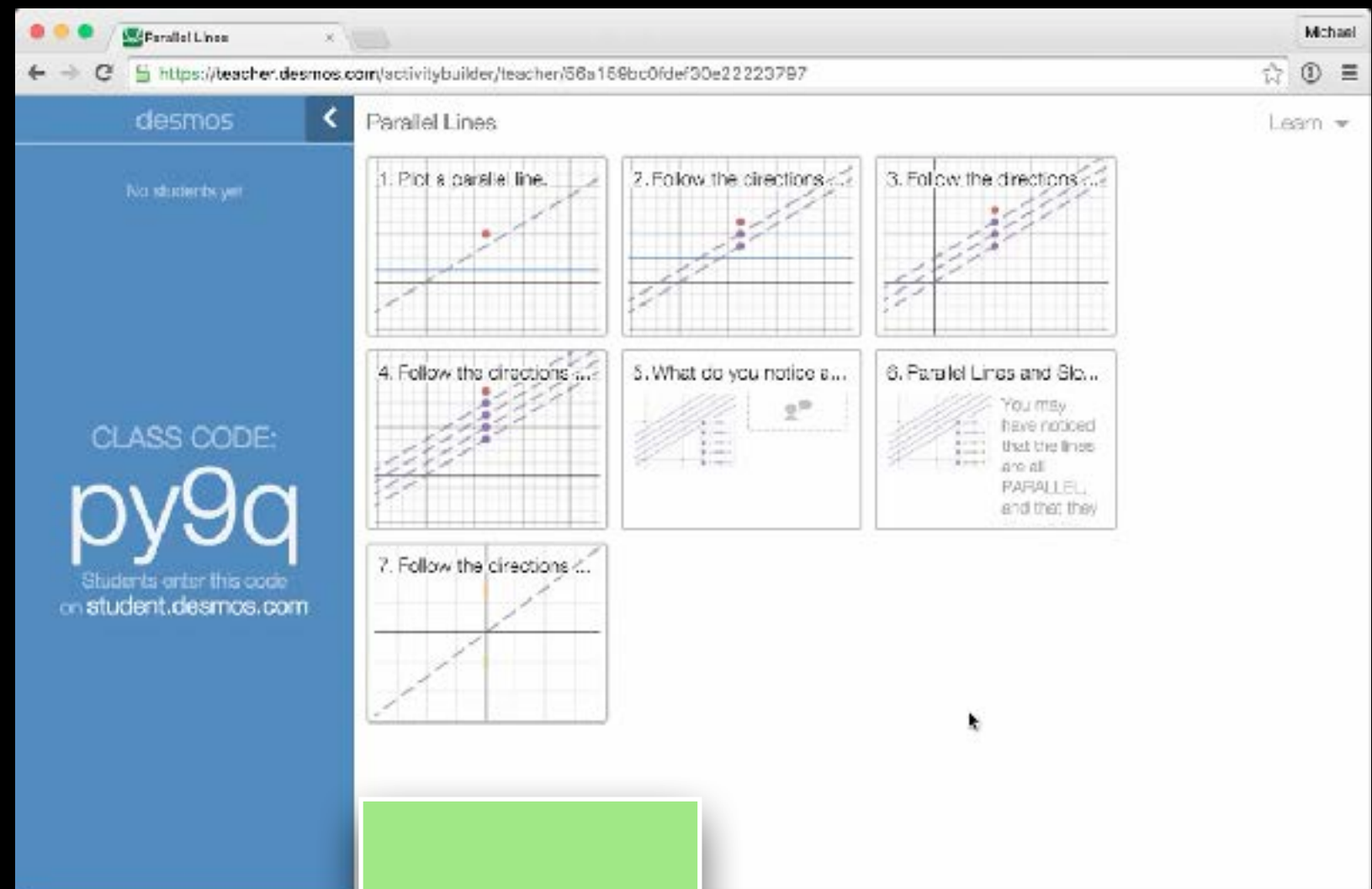


Realtime

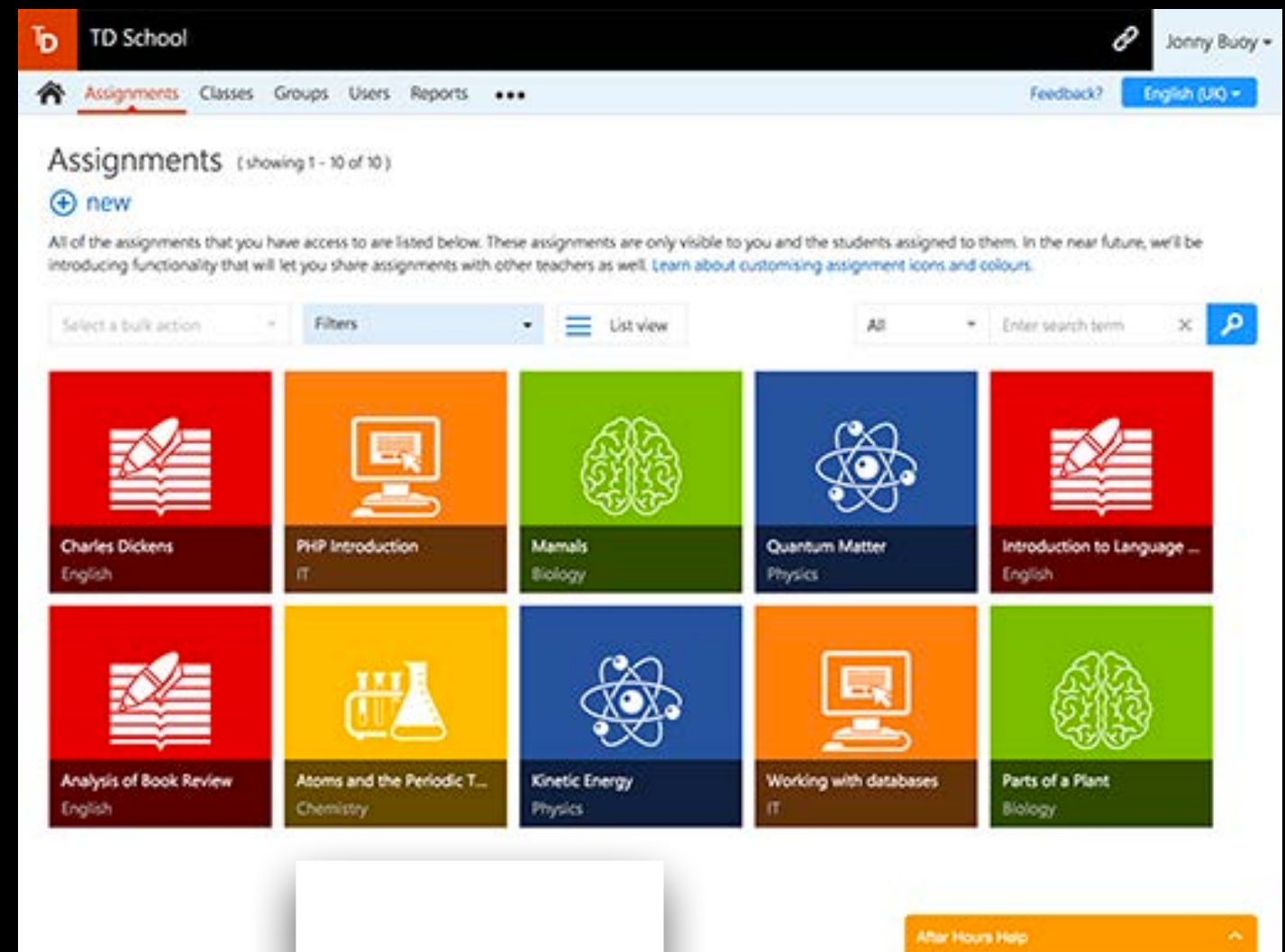
Data

Visualisation

Replay



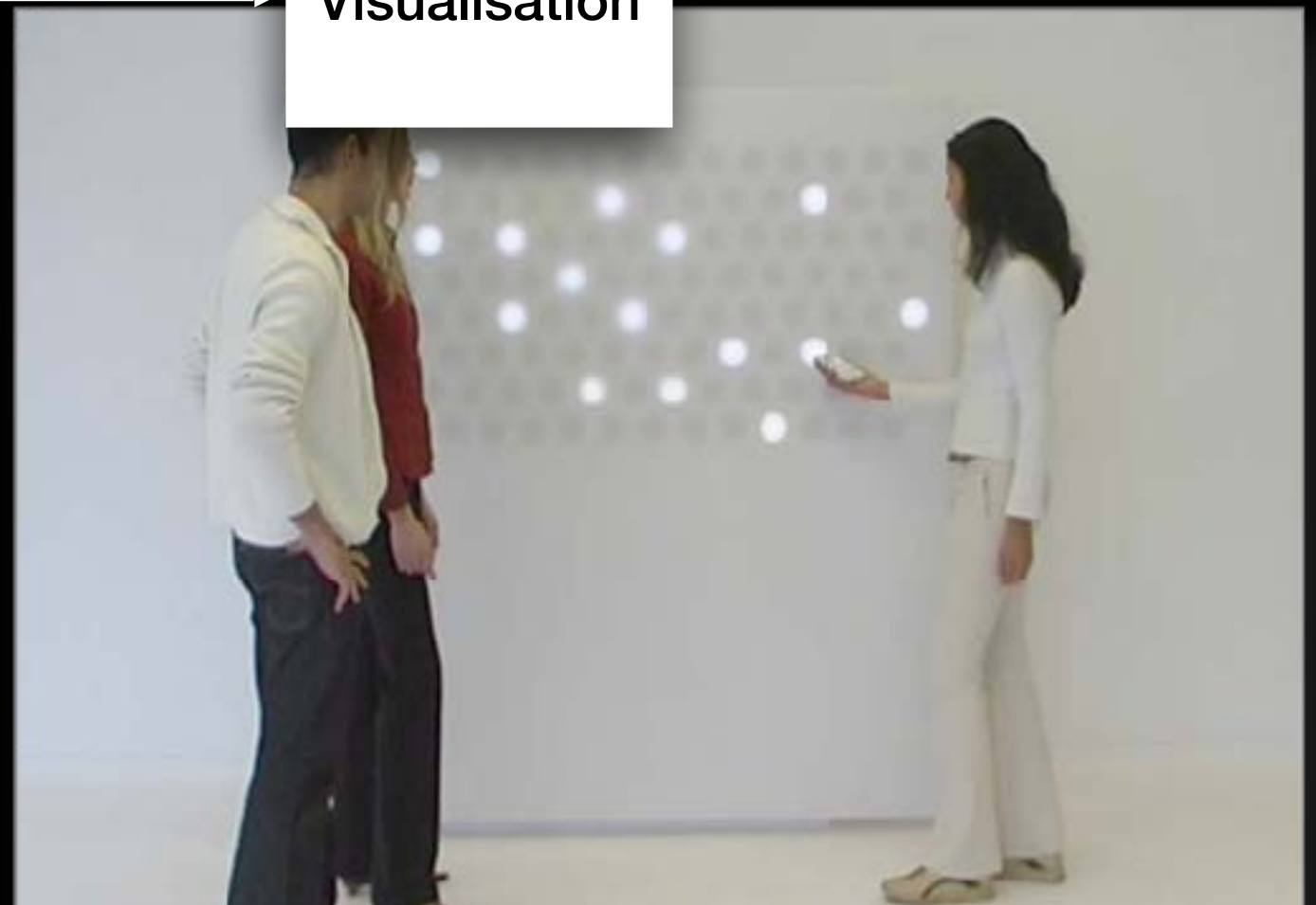
Focal



Data

Visualisation

Peripheral
(ambient)

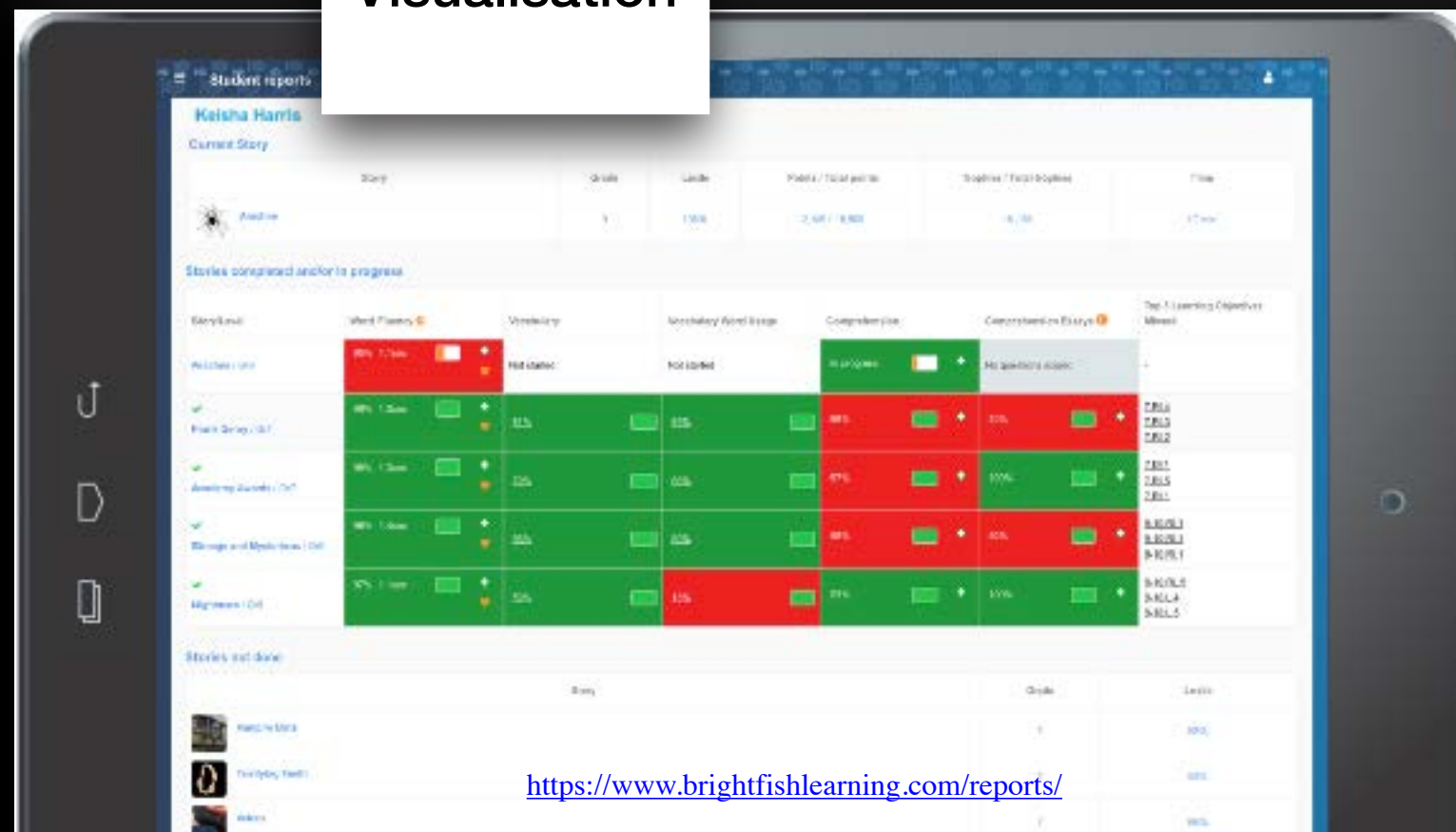


Public Dashboard

Data

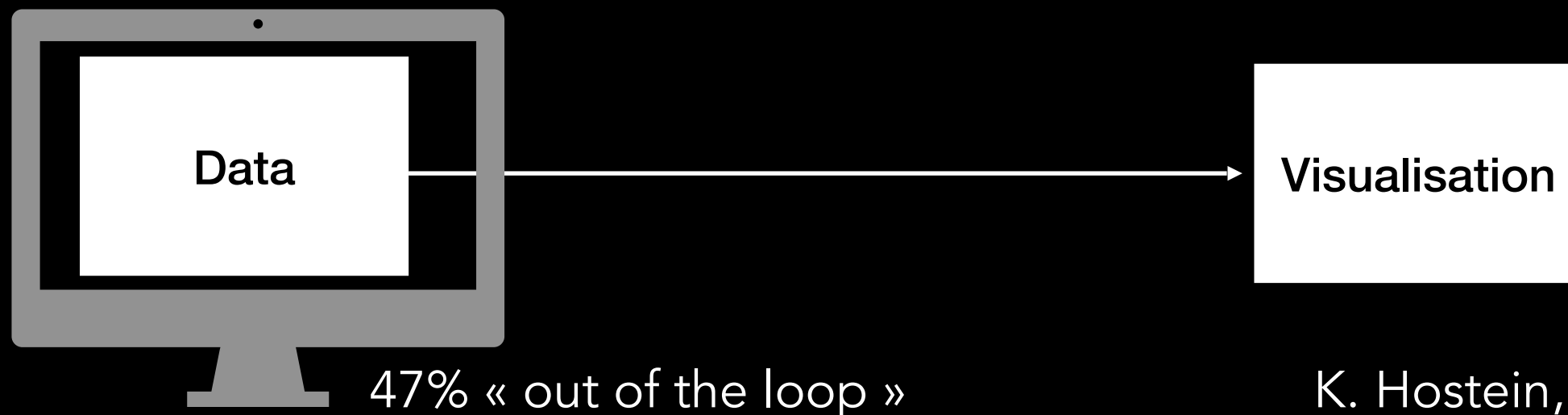
Visualisation

Private Dashboard

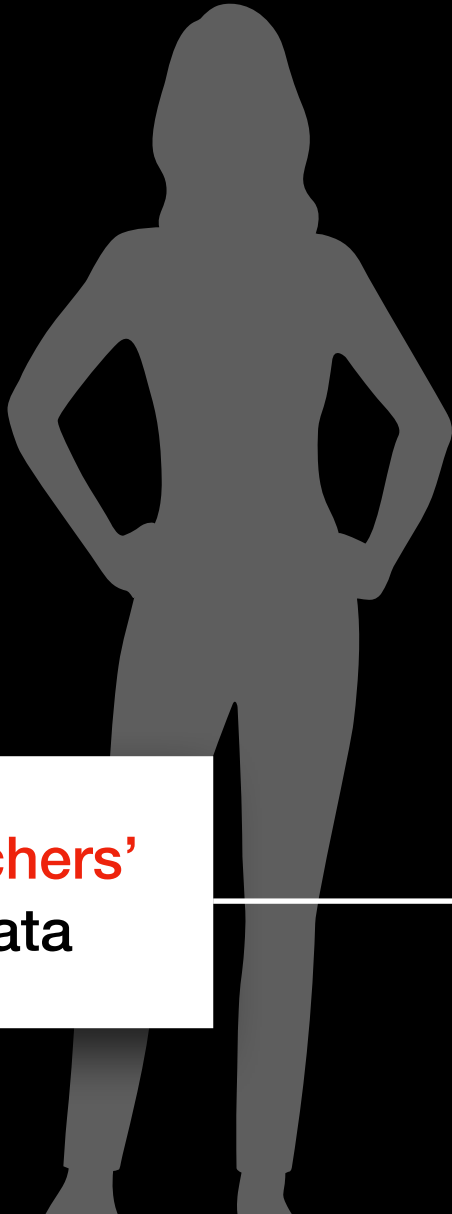


LUMILO

real-time, continuous assessment
for K-12 teachers



K. Hostein, V. Aleven & N. Rummel

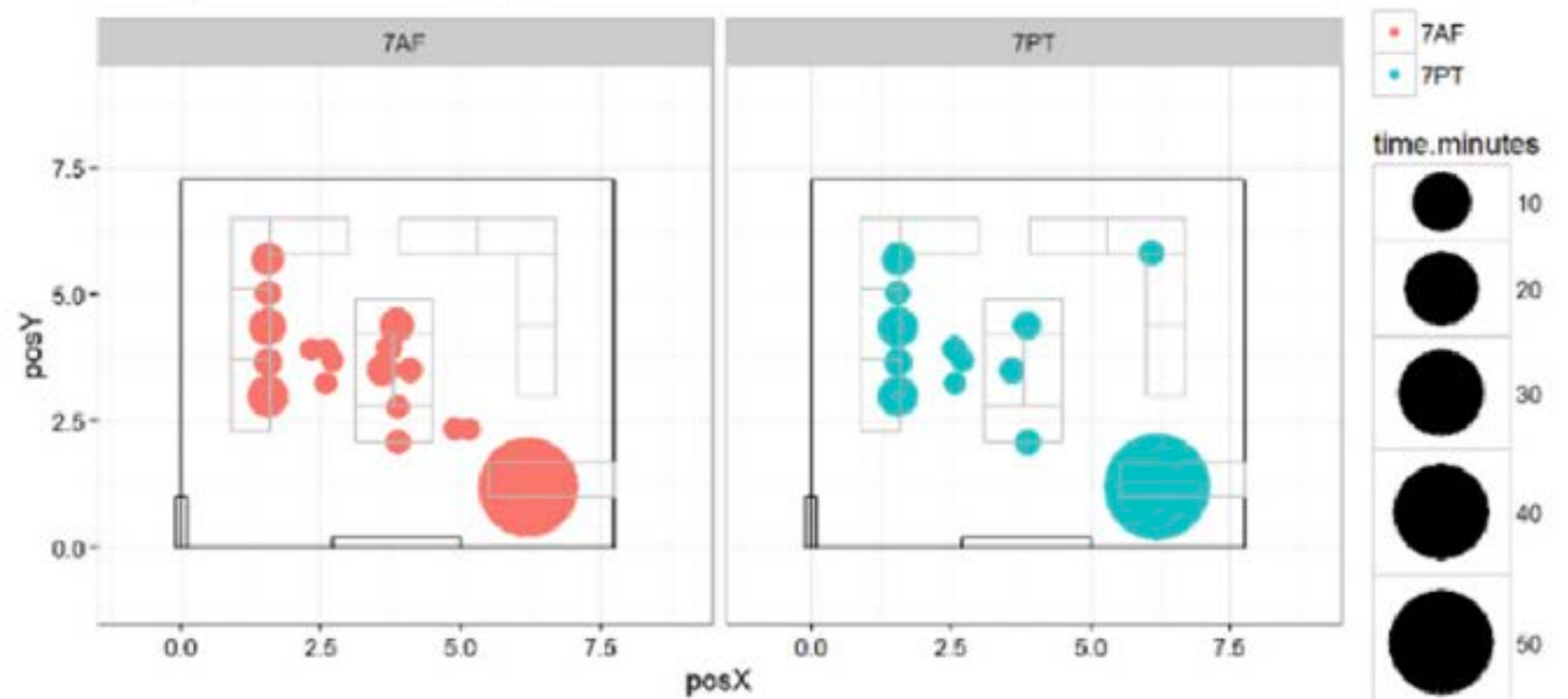


Teachers'
Data

Visualisation

Sarrade Isabelle, EPFL

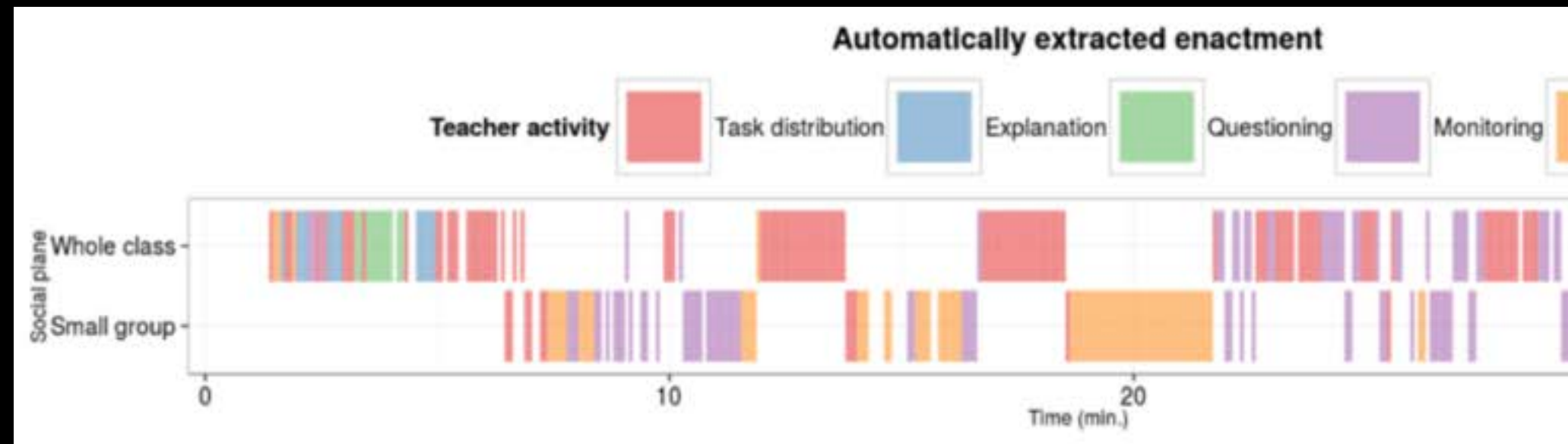
Overall position map per class



Prieto et al

Teachers'
Data

Visualisation



Teachers'
Data

Visualisation

Teachers'
Data

Visualisation



Orange



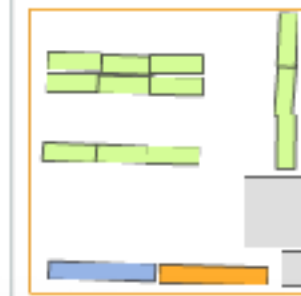
(1) monitoring

Comparaison

☒ Afficher les statistiques



Brute: 256m2
Brute stock.: 236m2
Nette: 30m2
Étagères: 12
Degré d'util.: 12.6%
Chariot: gerbeur
Temps/palette: 115s
Temps simul.: 0:13:49



Brute: 256m2
Brute stock.: 220m2
Nette: 36m2
Étagères: 12
Degré d'util.: 16.4%
Chariot: gerbeur
Temps/palette: 130s
Temps simul.: 0:23:40

(2) debriefing

Problem Solving

Lecture



Question

Please order a standard return 2nd class

Enter command
from Lausanne to Davos standard C2|re

Question

Please order a young return 2nd class ticket from Basel to Geneve without bike.

Your Ticket

From
Basel

To
Geneve

Travel
Return

Class

Fare

Bike

City

Basel

Geneve

Davos

Lausanne

Zurich

Fribourg

Neuchatel

Travel

Fare

Class

Bike

One-way

Return

Standard

Young

Half-fare

Question

Please order a standard return 2nd class ticket

From:
Fribourg

To:
Zurich

Travel:
Return

Fare:
Standard

Class:
2nd

Bike:
No

Question

Please order a standard return 2nd class ticket from Basel to Zurich with a bike.



From
Basel

To
Zurich

Travel

Fare

Class

Bike

☐ One-way

☐ Return

☐ Standard

☐ Young

☐ Half-fare

☐ 1st

☐ 2nd

☐ Yes

☐ No

⌚ :24

⌚ :36

HELP

BUY

Please select the interfaces and rank them with 1 being the best and 4 being the worst. Please justify your ranking.



If you rarely buy a train ticket rank the interfaces in the order that you would most prefer them.

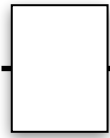
Ryan's List

1	Drag and Drop	↓	↑
2	Command	↓	↑
3	Form	↓	↑
4	Map	↓	↑

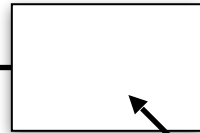
The drag and drop is easiest to see all of the options.

Submit

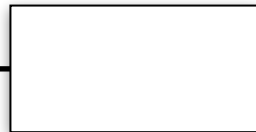
Class



Team



Solo



Chat (group/A1)

Group Chat

I had really liked the drag and drop because everything was so visible

Jenny

It was so slow though. All of the movement took forever

Ryan

What about the form? Everything was still visible but relatively fast with familiar interactions.

Jenny

yeah, ok

Jenny

What after that?

Group Preference (group/A1)

You and your partner must have the same ranking to submit.

Rank the interfaces in the order that you would most prefer them.

Ryan's List

Jenny's List

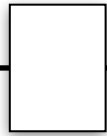
1 Form	1 Form	↓ ↑
2 Drag and Drop	2 Drag and Drop	↓ ↑
3 Command	3 Command	↓ ↑
4 Map	4 Map	↓ ↑

The form showed all the choices but was still pretty fast.

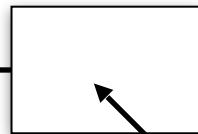
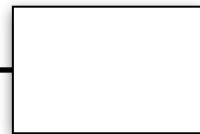
Submit

Socio-cognitive conflict

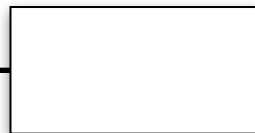
Class



Team



Solo



Arguing with data

Chat (group/alone)

Group Chat

Friendly robot
Hello Guys :) I <3 CHILIans

Friendly robot
Ryan ranked the interfaces in the following order: Command, Drag and Drop, Form, Map, with the justification "The command is fastest once you have practice."

Group Preference with Data (group/alone)

You and your partner must have the same ranking to submit.

Rank the interfaces in the order that you would most prefer them.

Ryan's List

At rank 1, add item:

Train Data (group/alone)

STATS

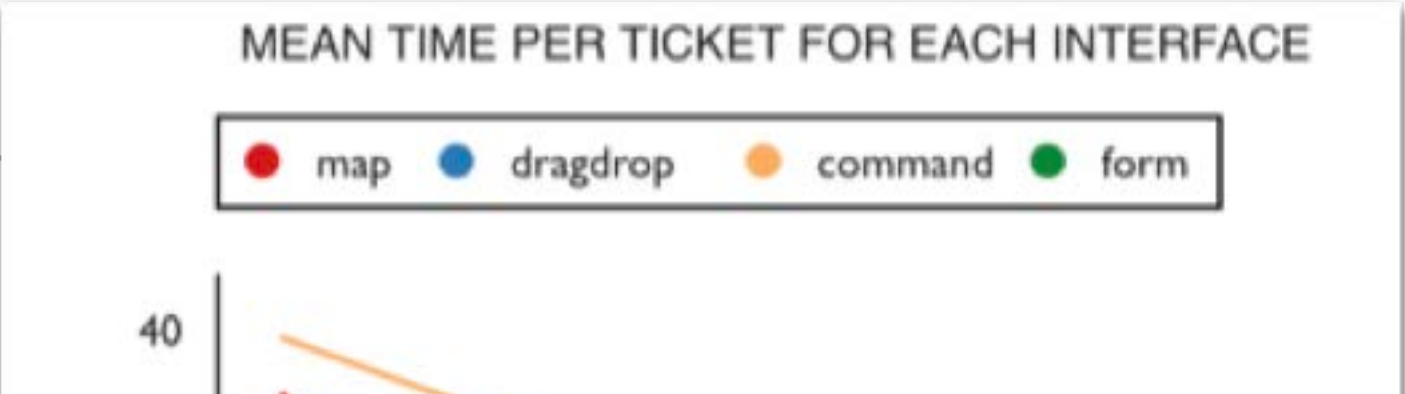
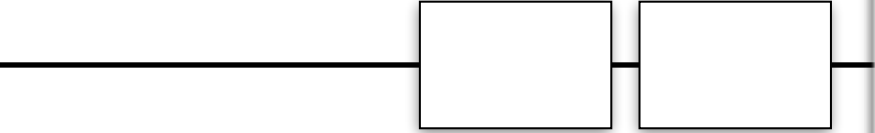
MEAN TIME PER TICKET FOR EACH INTERFACE

Trial	map	dragdrop	command	form
Trial 1	38	28	40	35
Trial 2	10	24	32	10
Trial 3	10	10	30	10

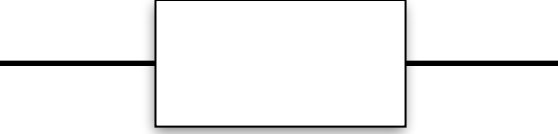
Class



Team



Solo



	(1) Connaissances sémantiques liées à la tâche	(2) Connaissances sémantiques liées à la transposition informatique de la tâche	(3) Connaissances syntaxiques, arbitraires
NOVICES	✓		
INTERMITTENTS	✓	✓	
EXPERTS	✓	✓	✓

Debriefing

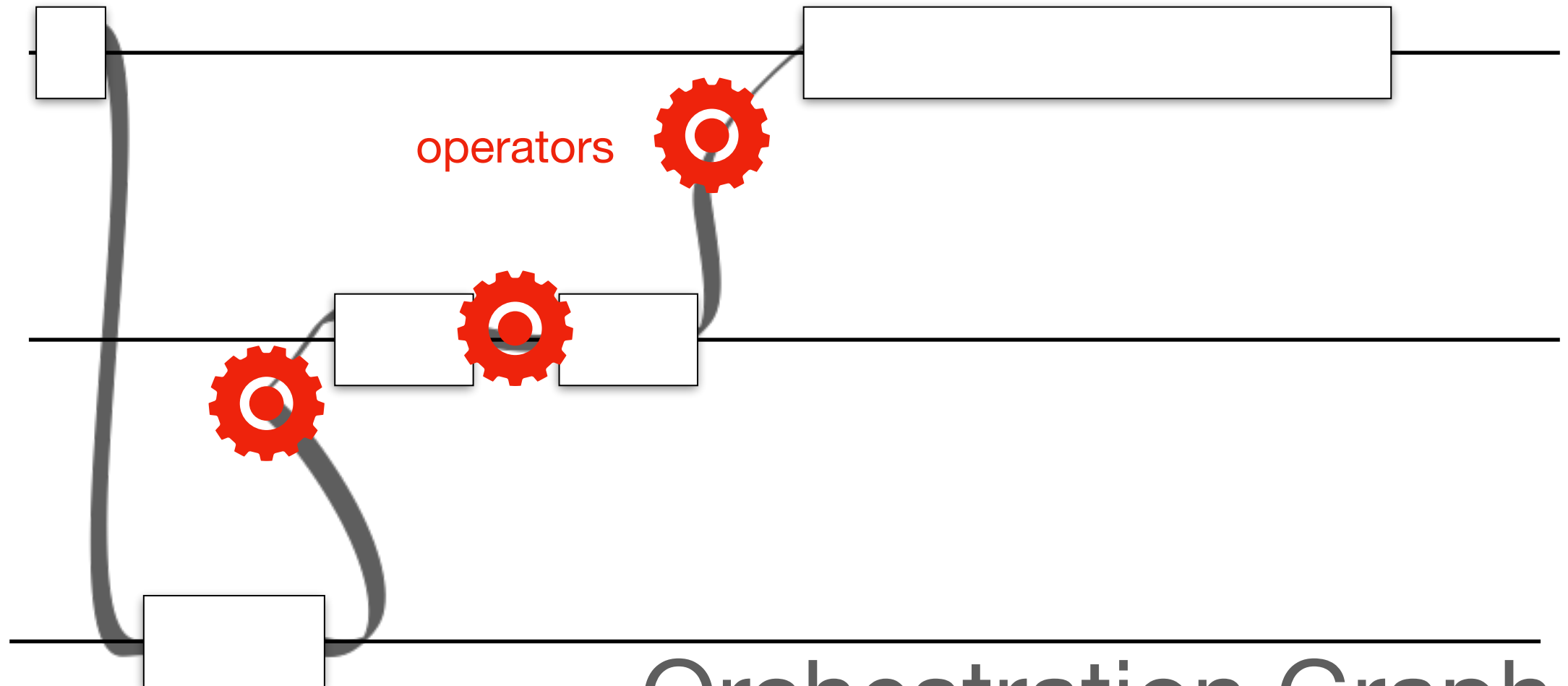
Class

Team

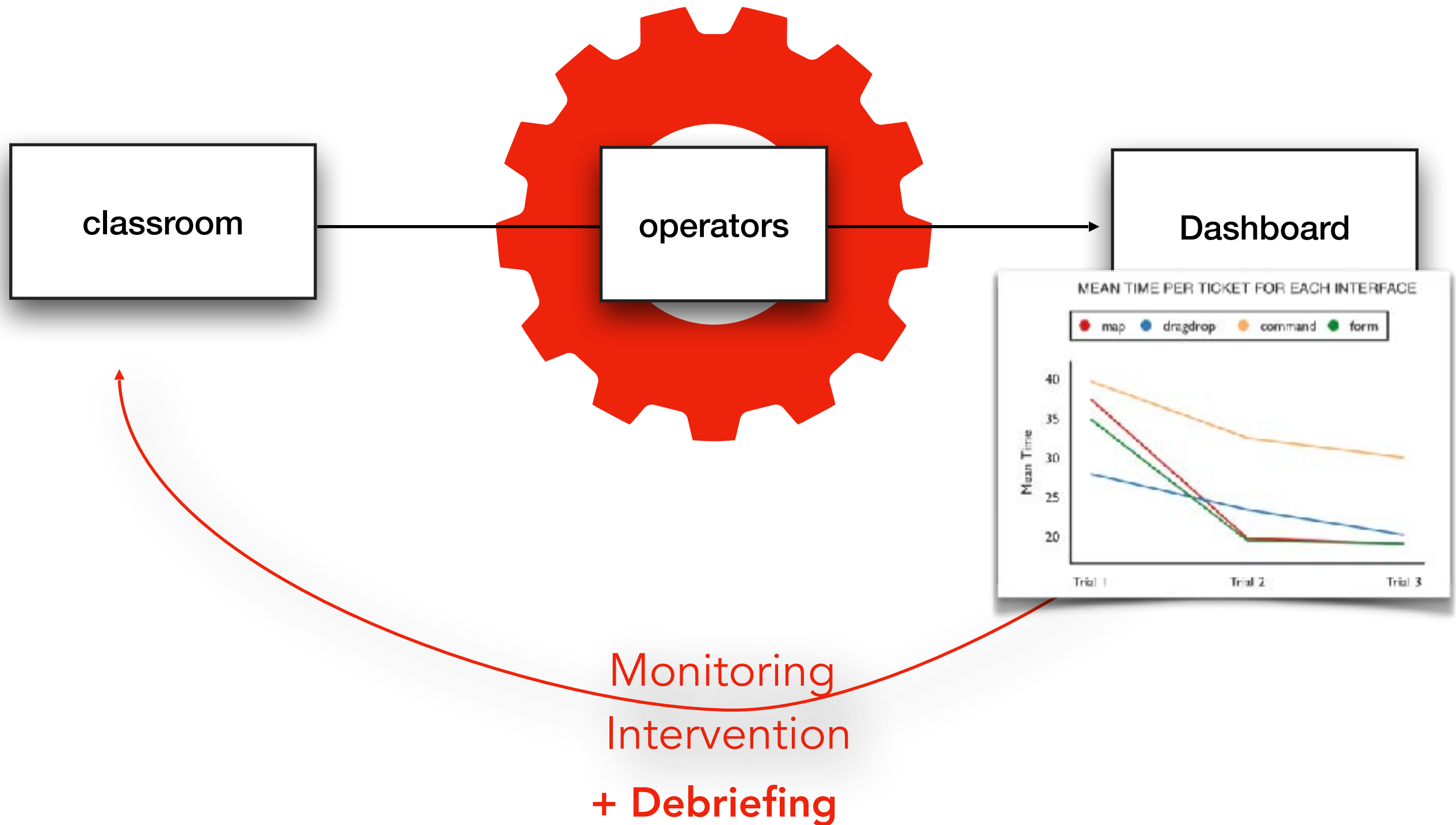
Solo

operators

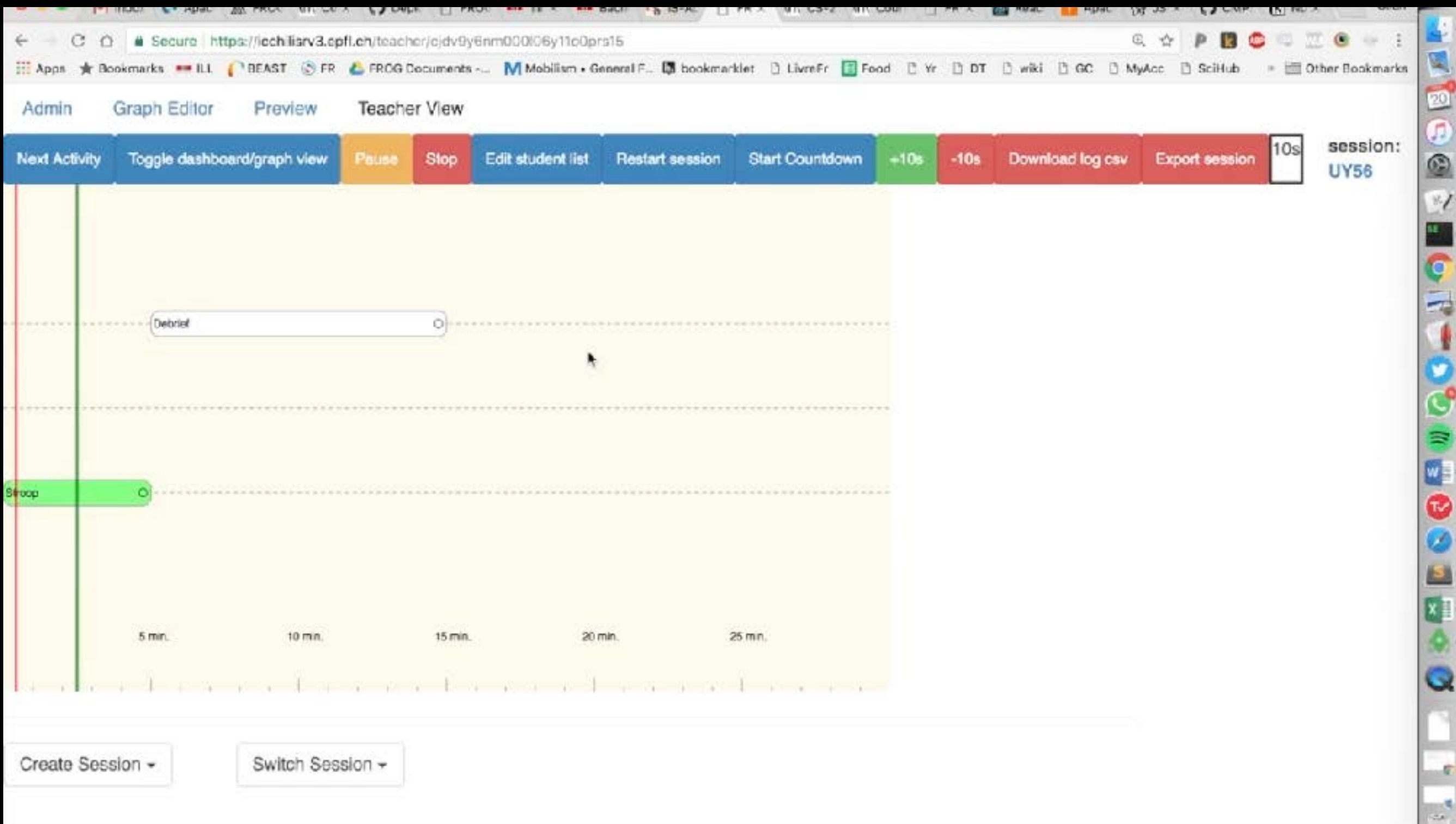
Orchestration Graph



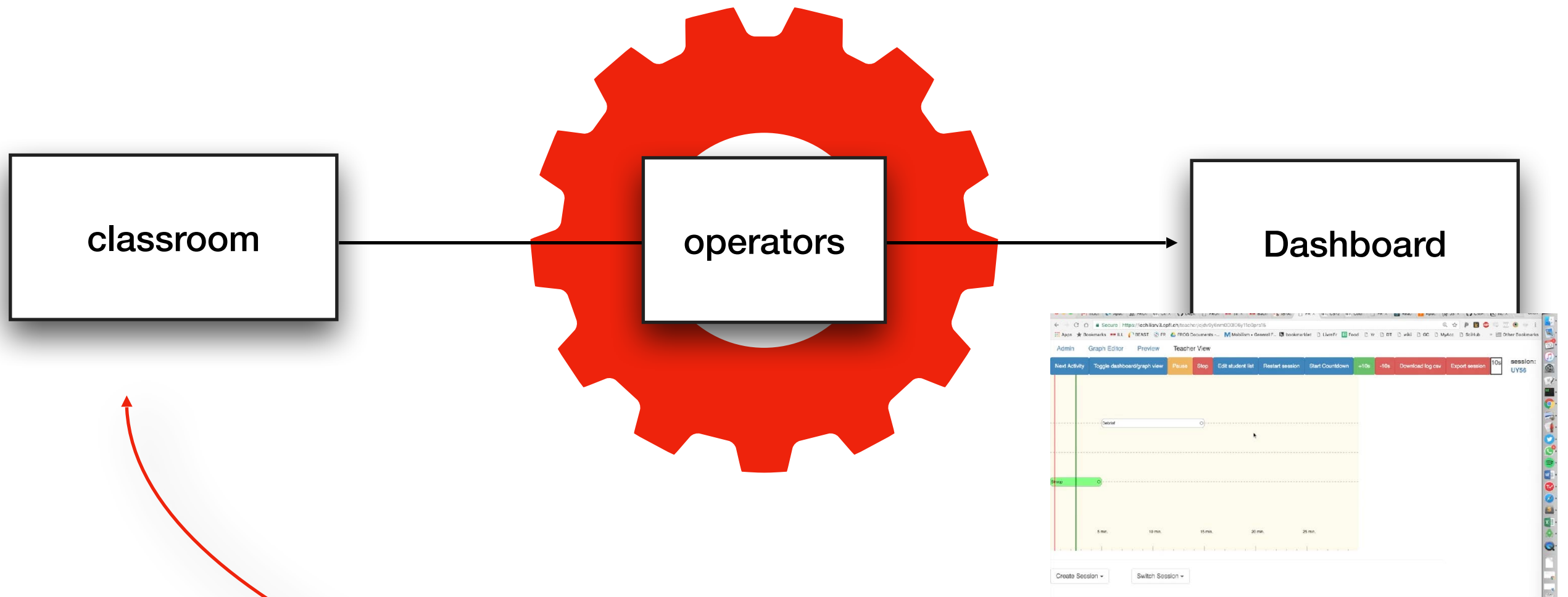
Aggregation, comparison, visualisation,....



Timing

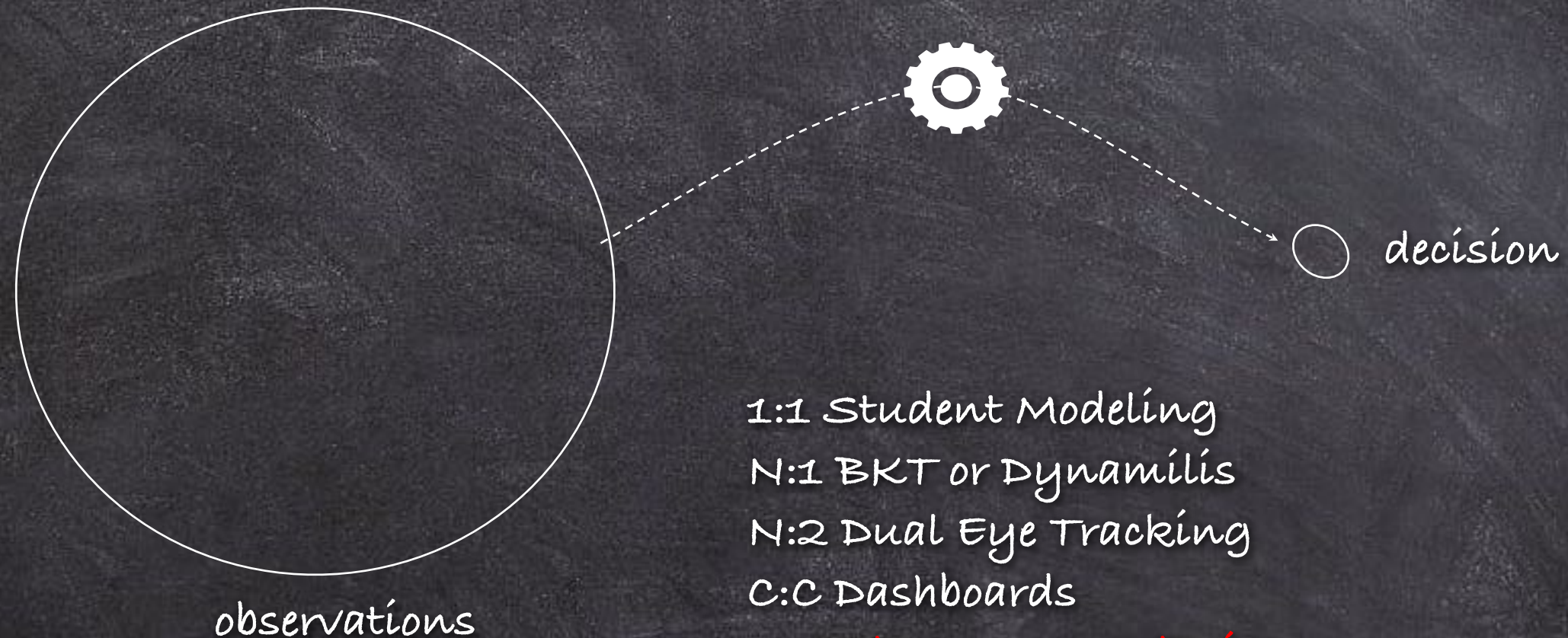


Aggregation, comparison, visualisation,....



Monitoring
Intervention
+ Debriefing
+ Timing

Learning Analytics Ratios



1:1 Student Modeling
N:1 BKT or Dynamilis
N:2 Dual Eye Tracking
C:C Dashboards
N:C Classroom analytics
N:N Social Network Analysis
N:S Training needs Analysis

$$2 < C < N < P$$