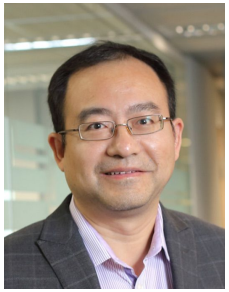
Interactive Visual Analytics for **Personalized** Online Learning

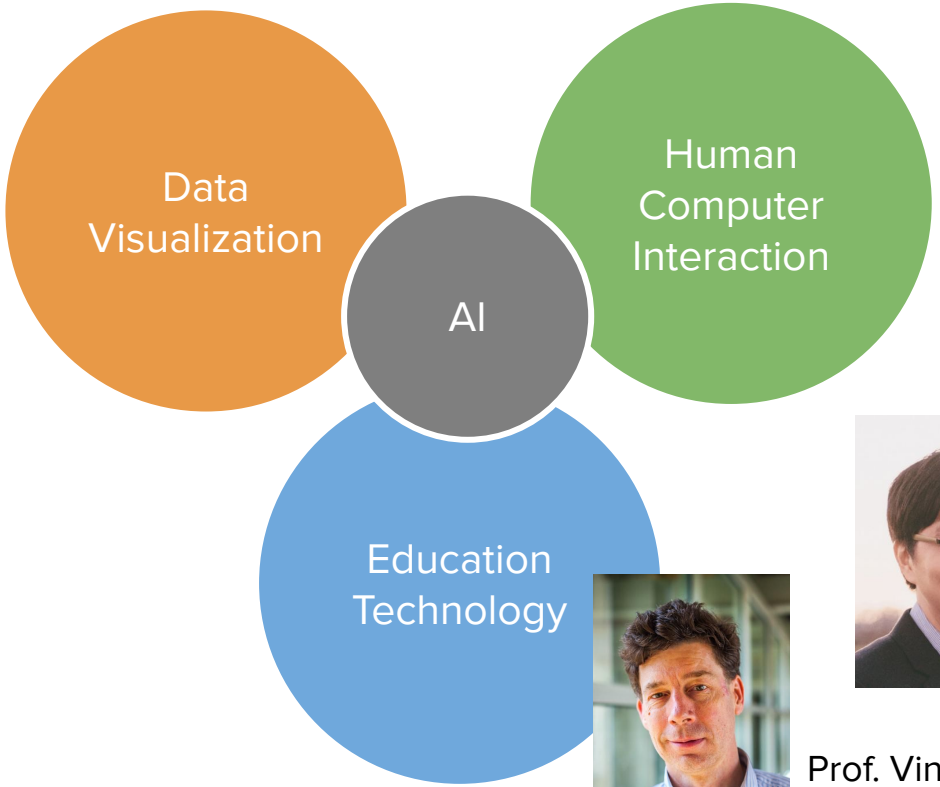
Meng Xia

07-01-2022

Research Background



Prof. Huamin Qu
(HKUST)



Prof. Xiaojuan Ma
(HKUST)

Prof. Juho Kim
(KAIST)

Prof. Vincent Aleven
(CMU)

What is Visual Analytics?

Visual analytics provides **visual representations** of datasets and interactive technologies to **augment** human's ability in finding **insights** in **data**

Input: data

Output: interactive visualizations

Goal: **augmenting** human's ability in finding **insights** in data

Why Visual Representation?

Anscombes quarte

I		II		III		IV	
x	*y*	*x*	*y*	*x*	*y*	*x*	*y*
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Table 1: Anscombe’s quartet: four different datasets

Why Visual Representation?

Anscombes quarte

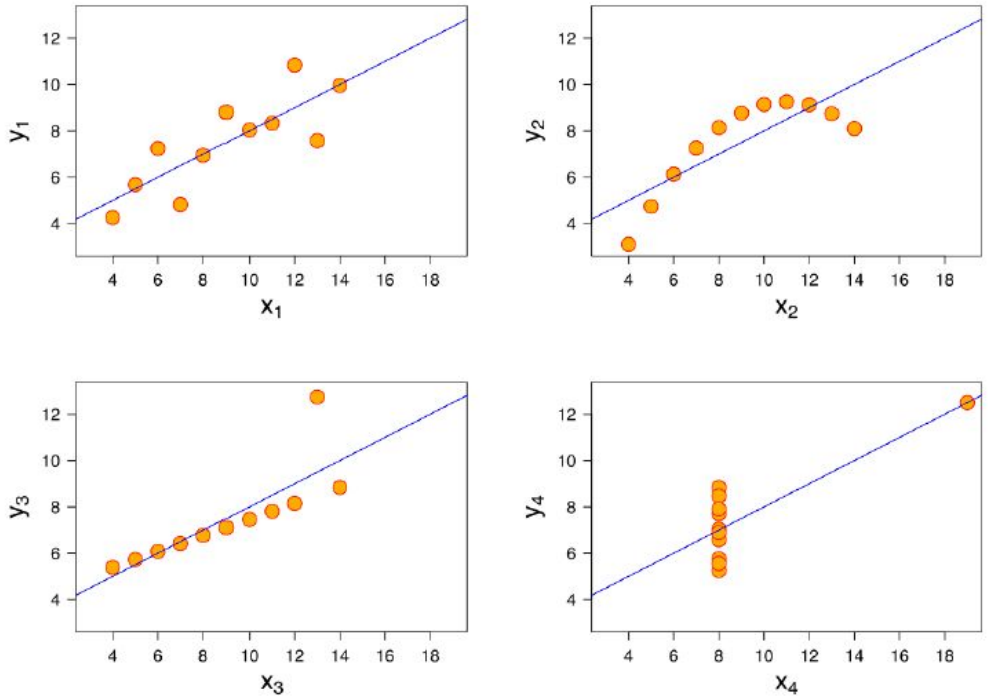
I		II		III		IV	
x	*y*	*x*	*y*	*x*	*y*	*x*	*y*
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Table 1: Anscombe's quartet: four different datasets

Property (in each set)	Value
Mean of x	9.0
Variance of x	10.0
Mean of y	7.50
Variance of y	3.75
Correlation between x and y	0.898
Linear regression line	$y = 0.5x + 3.0$

Table 2: Same statistics in Anscombe's quartet

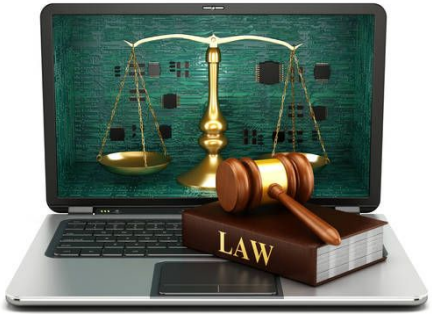
Why Visual Representation?

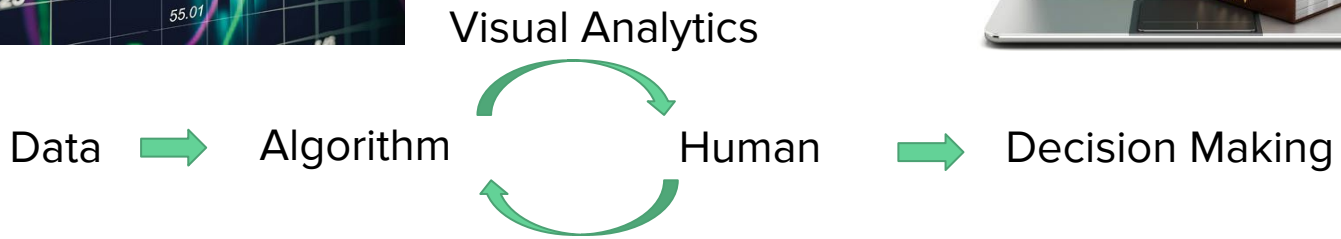


- Complement cognition with perception

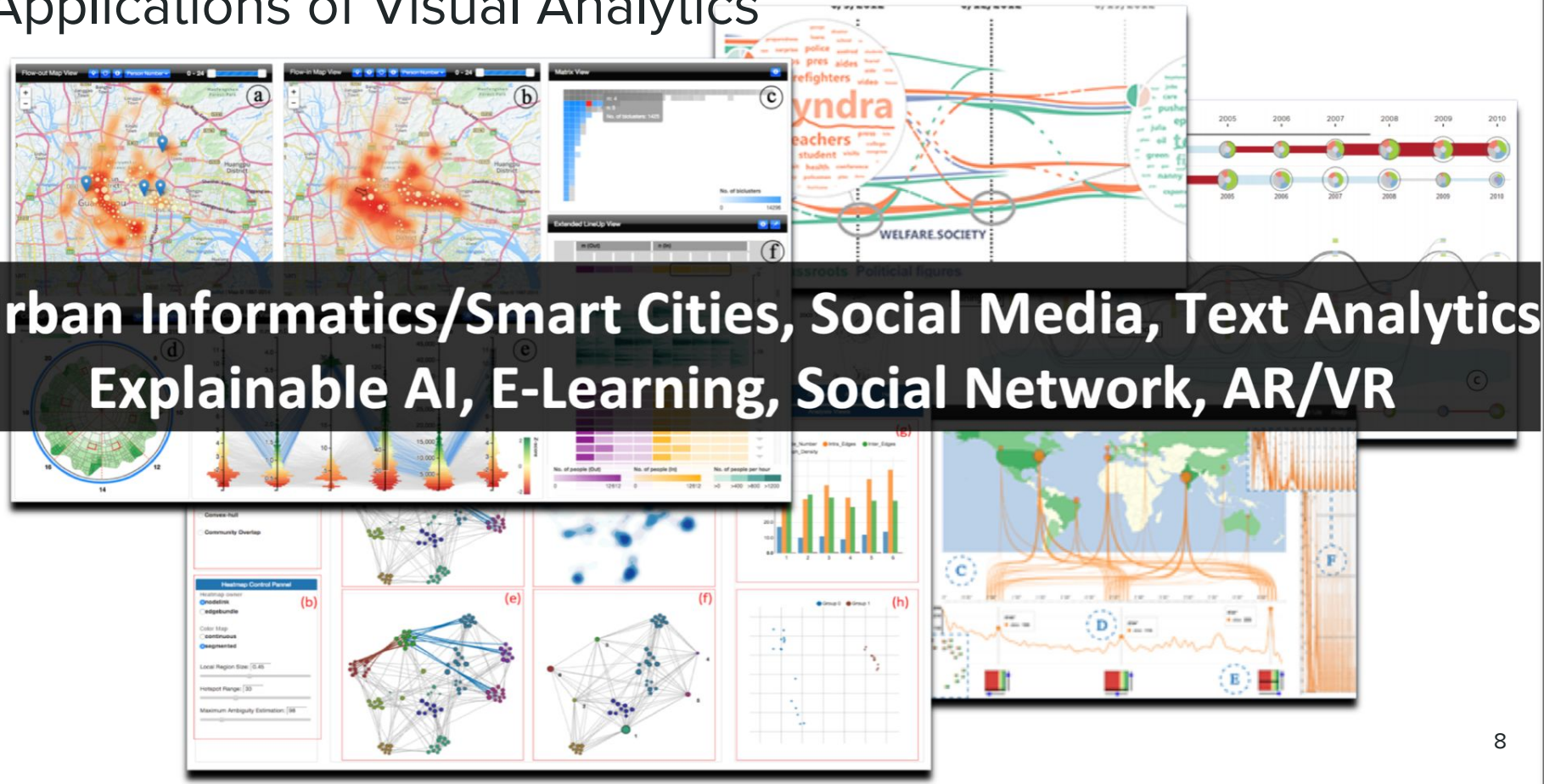
Why visual analytics? Keep Human in the Loop

Don’t need vis when fully automatic solution exists and is trusted. However, when there isn’t, visual analytics can help.

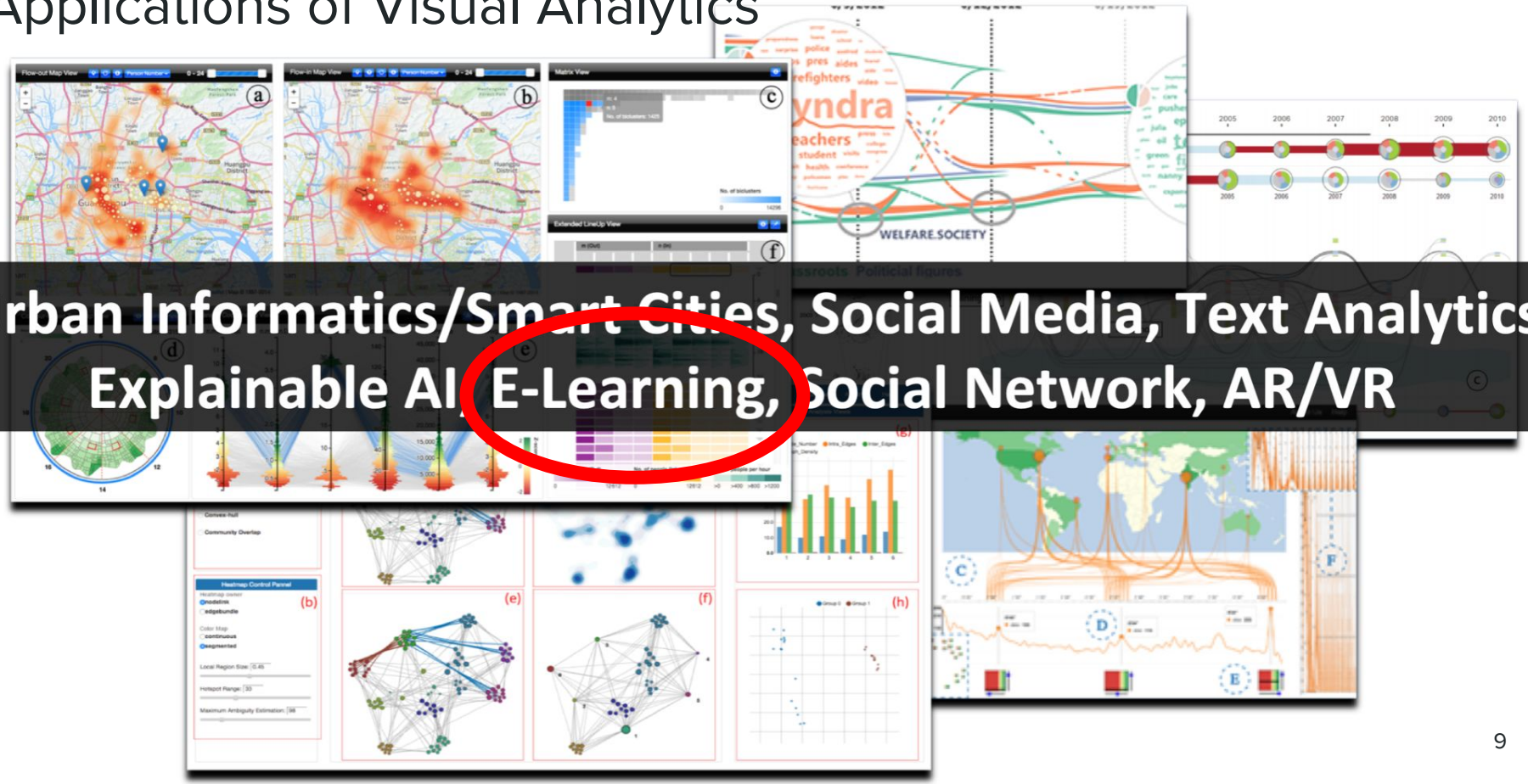




Applications of Visual Analytics



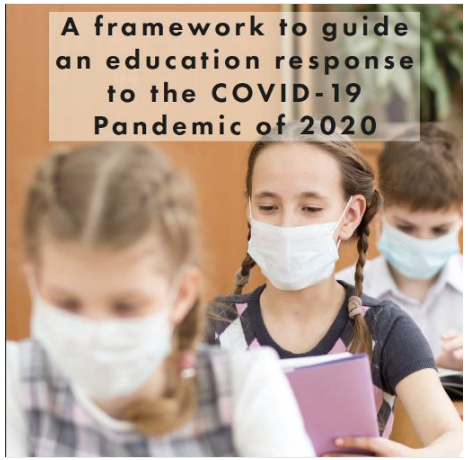
Applications of Visual Analytics

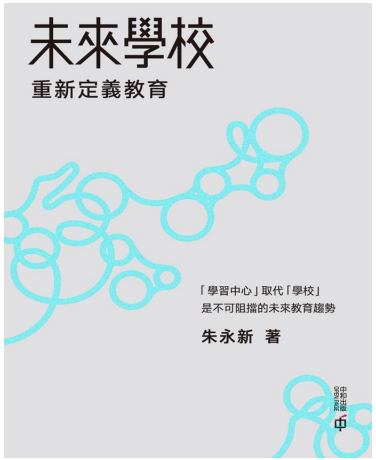


E-learning/Online Learning is Important

=> Flexible learning location.

=> Learning at Scale.





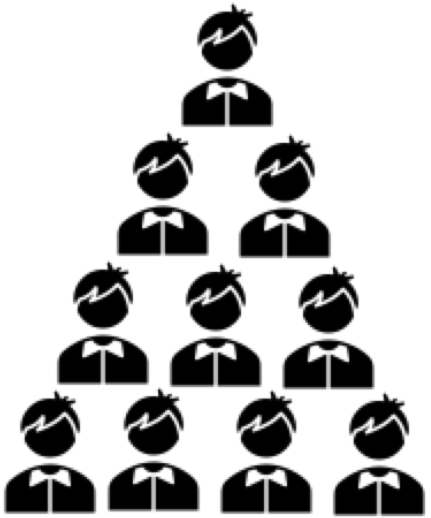
*A Framework to Guide and Education Response to the COVID-19 Pandemic by **OECD (Organization for Economic Co-operation and Development), 2020***

Challenges in Online Learning

=> Flexible learning location.

=> Learning at Scale.

=> Personalized Learning?



Personalized?

Powered by Learning data

Type	Online learning platforms	Examples	Learning Data
Self-Learning	Problem-based: Intelligent tutoring system/Test and quiz systems	Algebra Tutor, SmartTutor/LeetCode, Uva	Problem-solving data
	ill-defined tasks: Online forums, Q/A systems	Reddit/StackOverflow	Online forum data
	Video-based: Learning management system	Canvas, Moodle, Coursera, EdX, Udacity	Video watching data
Live Lessons	Online tutoring platforms	Cambly, Preply, italki, Zoom	Video and audio communications

Personalized Online Learning

Educators (design loop)

Designing
Personlized learning materials and instructions

Learners (learning loop)

Reflecting
Regulating learning behaviors

Planning
Customizing learning goals

Learning
Personalizing learning activities

Understanding and Analyzing
Different learners cognitive and non-cognitive behaviors

Data

Personalized Online Learning

Educators (design loop)

Redesigning
Personlized learning materials and instructions

Learners (learning loop)

Reflecting
Regulating learning behaviors

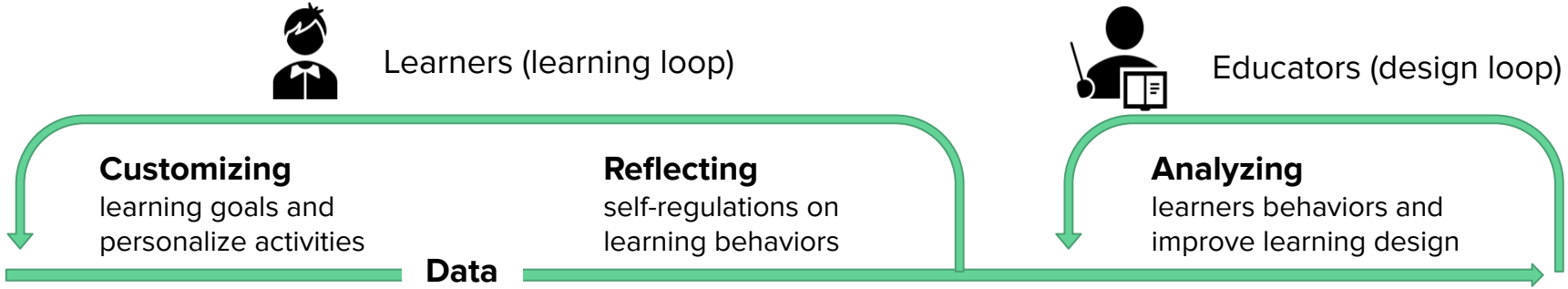
Planning
Customizing learning goals

Performing
Personalizing learning activities

Understanding and Analyzing
Different learners cognitive and non-cognitive behaviors

Data

Why Visual Analytics in Personalized Online Learning?



- Learning motivation, status, behaviors can hardly be defined and modelled using algorithms.
- Learning is a high-risk task that needs careful decision making.

Challenges

Learners (learning loop)

Educators (design loop)

Customizing
learning goals and personalize activities

Reflecting
self-regulations on learning behaviors

Analyzing
learners behaviors and improve learning design

Data

- Large heterogeneous data
- Limited expertise and time in data analysis
- Not enough guidance and explanations
- No guarantee of data quality

- Learning motivation, status, behaviors can hardly be defined and modelled using algorithms.
- Learning is a high-risk task that needs careful decision making.

Challenges

Learners (learning loop)

Educators (design loop)

Customizing
learning goals and personalize activities

Reflecting
self-regulations on learning behaviors

Analyzing
learners behaviors and improve learning design

Data

- Large heterogeneous data
- Limited expertise and time in data analysis
- Not enough guidance and explanations
- No guarantee of data quality

- Large heterogeneous data
- Limited expertise and time in data analysis
- Lack of motivation, consistent mental model, and actionable plans

- Learning motivation, status, behaviors can hardly be defined and modelled using algorithms.
- Learning is a high-risk task that needs careful decision making.

Challenges

Learners (learning loop)

Educators (design loop)

Customizing
learning goals and personalize activities

Reflecting
self-regulations on learning behaviors

Analyzing
learners behaviors and improve learning design

Data

- Large heterogeneous data
- Limited expertise and time in data analysis
- Not enough guidance and explanations
- No guarantee of data quality

- Large heterogeneous data
- Limited expertise and time in data analysis
- Lack of motivation, consistent mental model, and actionable plans

- Large heterogeneous data
- Limited expertise and time in data analysis
- No predefined model

- Learning motivation, status, behaviors can hardly be defined and modelled using algorithms.
- Learning is a high-risk task that needs careful decision making.

My works

Learners (learning loop)

Educators (design loop)

Customizing
learning goals and personalize activities

Reflecting
self-regulations on learning behaviors

Analyzing
learners behaviors and improve learning design

Data

- Large heterogeneous data
- Limited expertise and time in data analysis
- Not enough guidance and explanations
- No guarantee of data quality

- **AlgoSolve** (CHI 2022)
- **Persua** (CSCW 2021)
- **Peerlens** (CHI 2019)

- Large heterogeneous data
- Limited expertise and time in data analysis
- Lack of motivation, consistent mental model, and actionable plans

- **RLens** (L@S 2022)
- **“Game the system”** (L@S 2020)

- Large heterogeneous data
- Limited expertise and time in data analysis
- No predefined model

- **BlockLens** (L@S WIP 2022)
- **QLens** (TVCG 2021)
- **Predicting** (LAK 2020)
- **SeqDynamics** (EuroVis 2020)
- **K-12 Mathematics** (VIS 2019, **Best Poster Award**)

My works

Learners (learning loop)

Educators (design loop)

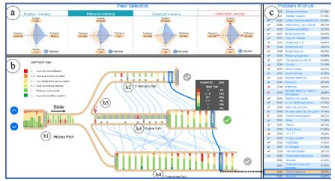
Customizing
learning goals and personalize activities

Reflecting
self-regulations on learning behaviors

Analyzing
learners behaviors and improve learning design

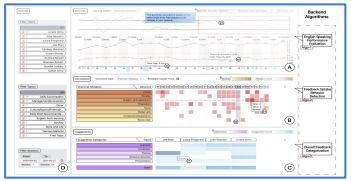
Data

- Large heterogeneous data
- Limited expertise and time in data analysis
- Not enough guidance and explanations
- No guarantee of data quality



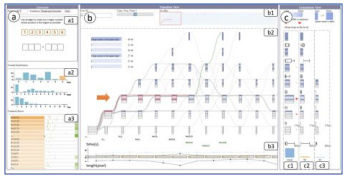
Peerlens (CHI 2019)

- Large heterogeneous data
- Limited expertise and time in data analysis
- Lack of motivation, consistent mental model, and actionable plans

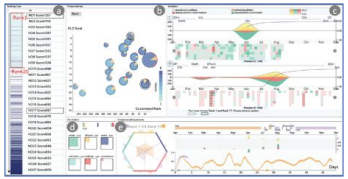


RLens (L@S 2022)

- Large heterogeneous data
- Limited expertise and time in data analysis
- No predefined model



QLens
(TVCG 2021)



SeqDynamics
(EuroVIS 2020)

My works

Learners (learning loop)

Educators (design loop)

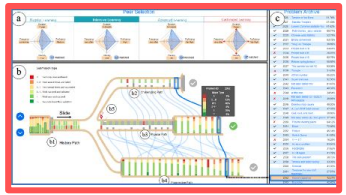
Customizing
learning goals and personalize activities

Reflecting
self-regulations on learning behaviors

Analyzing
learners behaviors and improve learning design

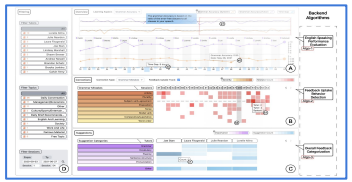
Data

- Large heterogeneous data
- Limited expertise and time in data analysis
- Not enough guidance and explanations
- No guarantee of data quality



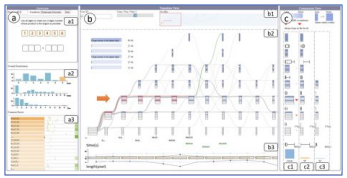
Peerlens (CHI 2019)

- Large heterogeneous data
- Limited expertise and time in data analysis
- Lack of motivation, consistent mental model, and actionable plans

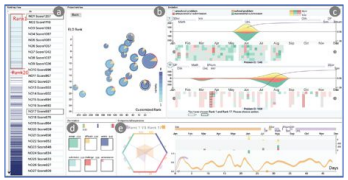


RLens (L@S 2022)

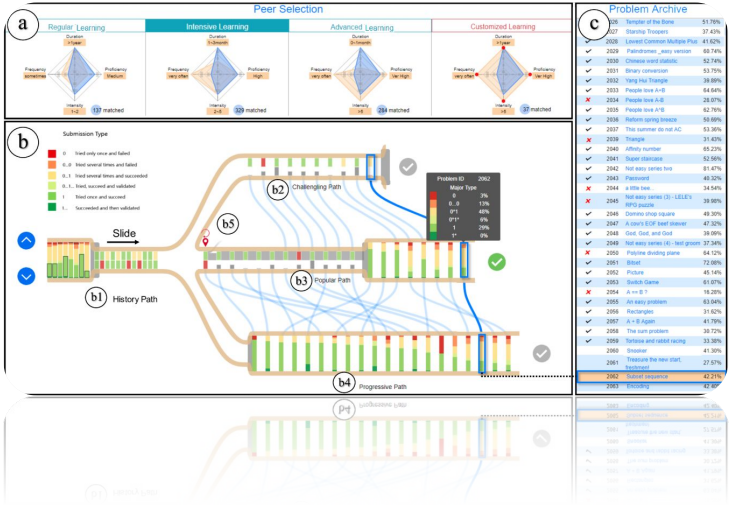
- Large heterogeneous data
- Limited expertise and time in data analysis
- No predefined model



QLens
(TVCG 2021)



SeqDynamics
(EuroVIS 2020)

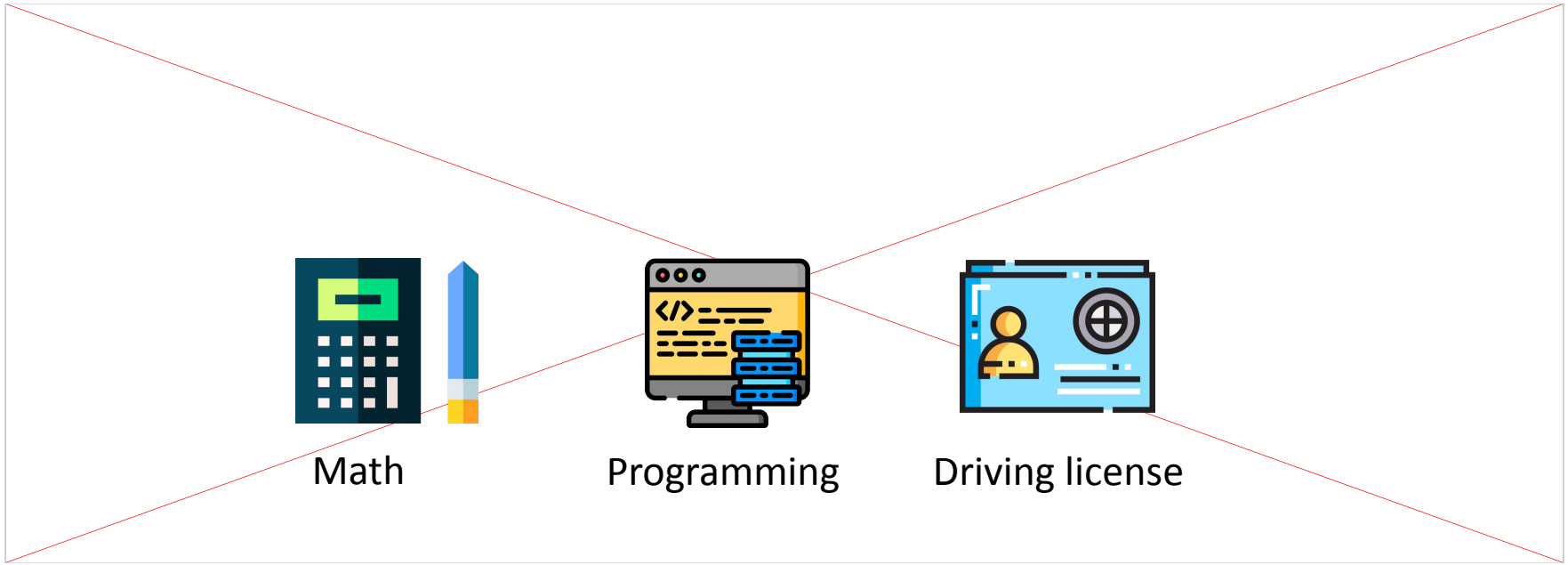


PeerLens: Peer-inspired Interactive Learning Path Planning in Online Question Pool

Meng Xia, Mingfei Sun, Huan Wei, Qing Chen, Yong Wang, Lei Shi, Huamin Qu, Xiaojuan Ma

CHI 2019

What is an online question pool?



https://help.blackboard.com/Learn/Instructor/Tests_Pools_Surveys/Reuse_Questions/Question_Pools_Banks

Features of question pools

Pro. ID	
1000	A + B Problem
1001	Sum Problem
1002	A + B Problem II
1003	Max Sum
1004	Let the Balloon Rise
1005	Number Sequence
1006	Tick and Tick
1007	Quoit Design
1008	Elevator
1009	FatMouse' Trade
1010	Tempter of the Bone
1011	Starship Troopers

- No pre-determined syllabus
- A lengthy list indexed by their problem IDs
- Hidden intents

- Different learning scenarios
- One learner's learning scenario may be changing

Difficulty: Determine an appropriate order in taking these online questions for their particular learning scenarios

Current situation

Programming question pools	Has recommendation?
AtCoder	NO
CodeChef	NO
CodeFights	NO
Codeforces	NO
Codewars	YES (Similar questions)
LeetCode	YES (Similar questions)
CodinGame	NO
Coderbyte	NO
CSAcademy	NO
HackerEarth	NO

Programming question pools	Has recommendation?
HackerRank	NO
Kattis	NO
uDebug	NO
OmegaUp	NO
Sphere Online Judge	NO
Topcoder	NO
Toph	NO
URI Online Judge	NO
UVa Online Judge	NO

Demand: planning personalized learning path in the context of existing list-based question pools

Related work: Educational Recommendation Techniques

Memory-based techniques

Continuously analyze current data *(Drachsler et al., 2008)*

- Content-based *(e.g., Chu et al., 2011)*, Collaborative Filtering *(e.g., Toledo et al., 2018)*, Hybrid approach *(e.g., Salehi et al., 2013)*

Lack of information

Model-based techniques

Utilize a large amount of historical data to model the learning process over time

- Deep learning models *(e.g., Piech et al., 2015)*, other models, such as Markov Chain *(e.g., Rajapakse and Ho, 2005; Sarukkai 2000; Huang et al., 2009)*

No explanation on the recommendations

A user-centered design process

- Four domain experts
  - Instruction designers who designed online question pools (E1, E2)
  - Online question pool users (S1, S2)

- Requirements gathering iteratively for three months

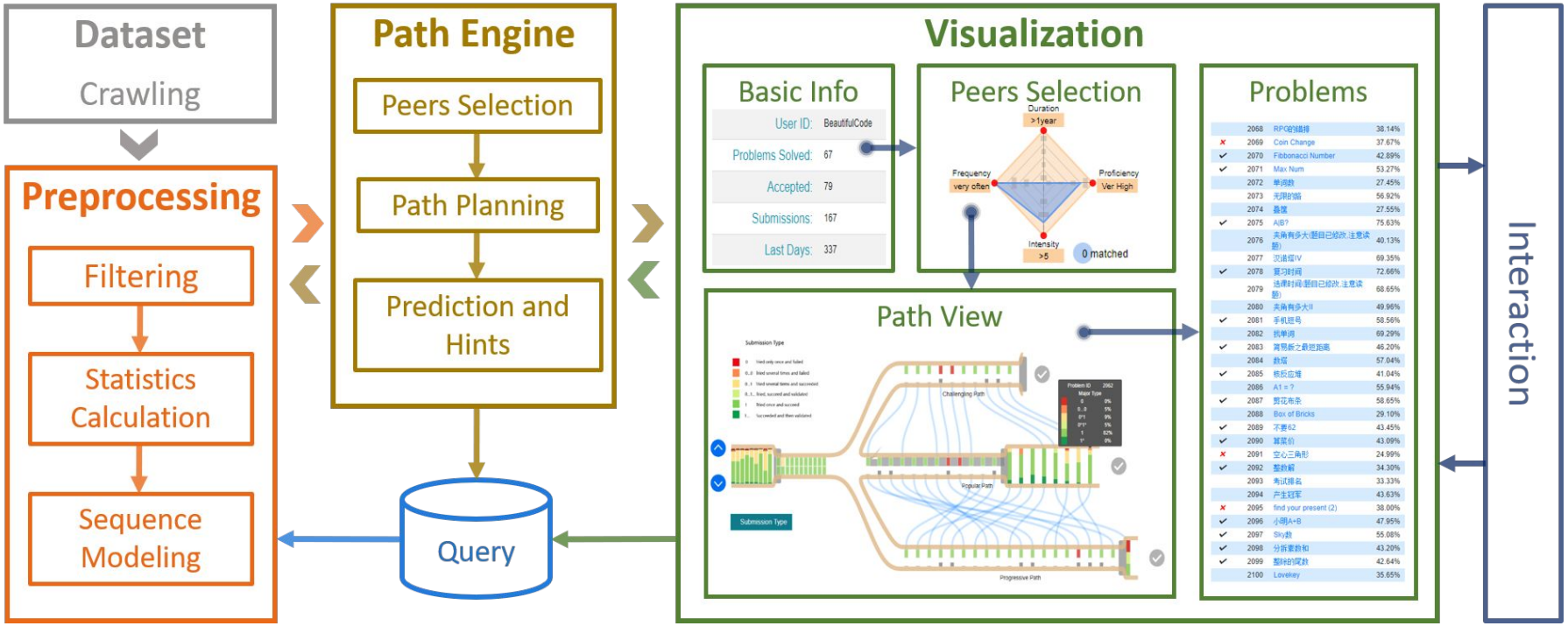
  R1: Find peers for a specific learning scenario.

  R2: Compare with peers' performance.

  R3: Offer flexible learning path suggestions with explanations.

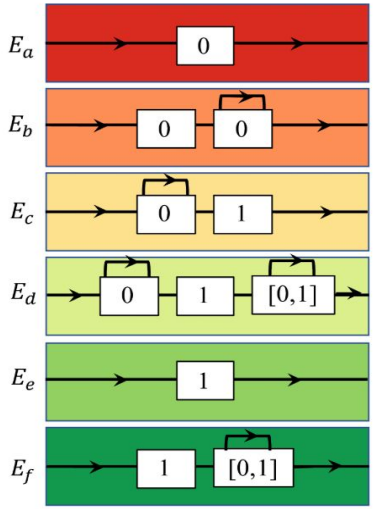
  R4: Provide convenient interaction and intuitive visual designs for learning path planning.

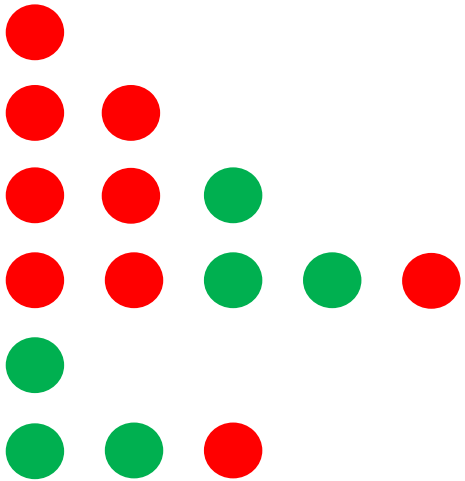
System overflow

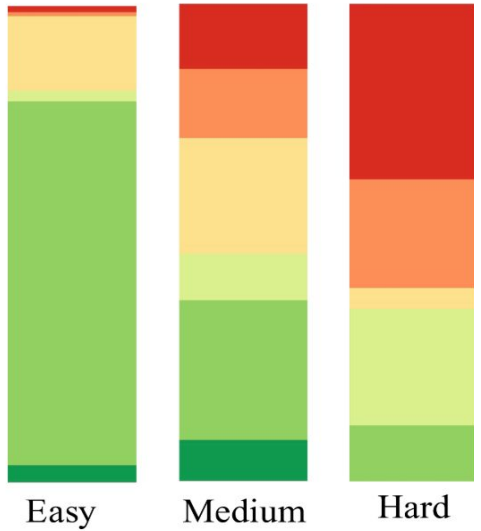


Path Planning Engine: Learning Path Modeling

Submission type: the way a user interacts with a problem.







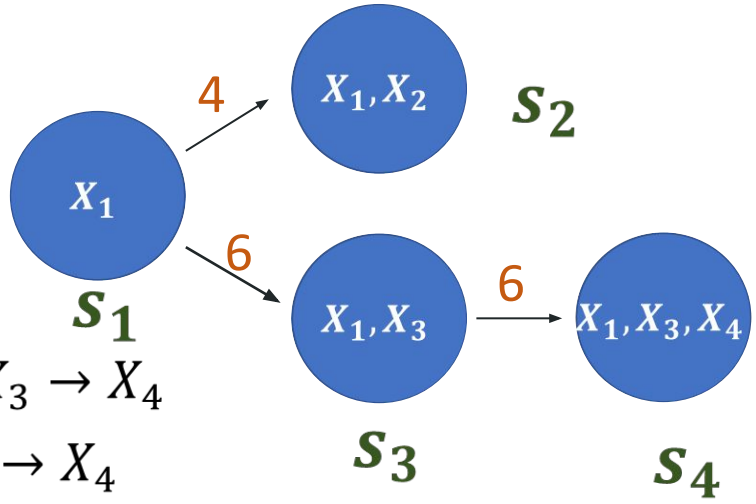
- Captures learners' knowledge proficiency
- Enables the inference of question difficulty level

incorrect correct

Path Planning Engine: Path Suggestion

A given peer path $[(X_{i_0}, E_{i_0}, t_{i_0}), \ldots, (X_{i_n}, E_{i_n}, t_{i_n})]$ corresponds to a state $s = \{X_{i_0}, X_{i_1}, \ldots, X_{i_n}\}$.

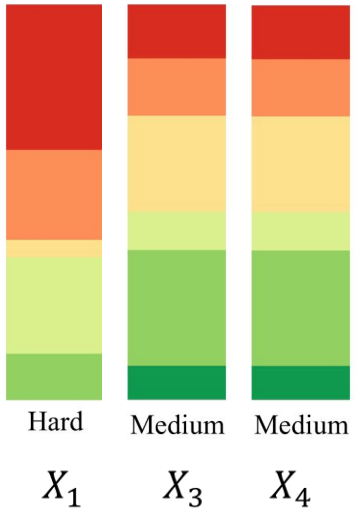
Markov Chain:

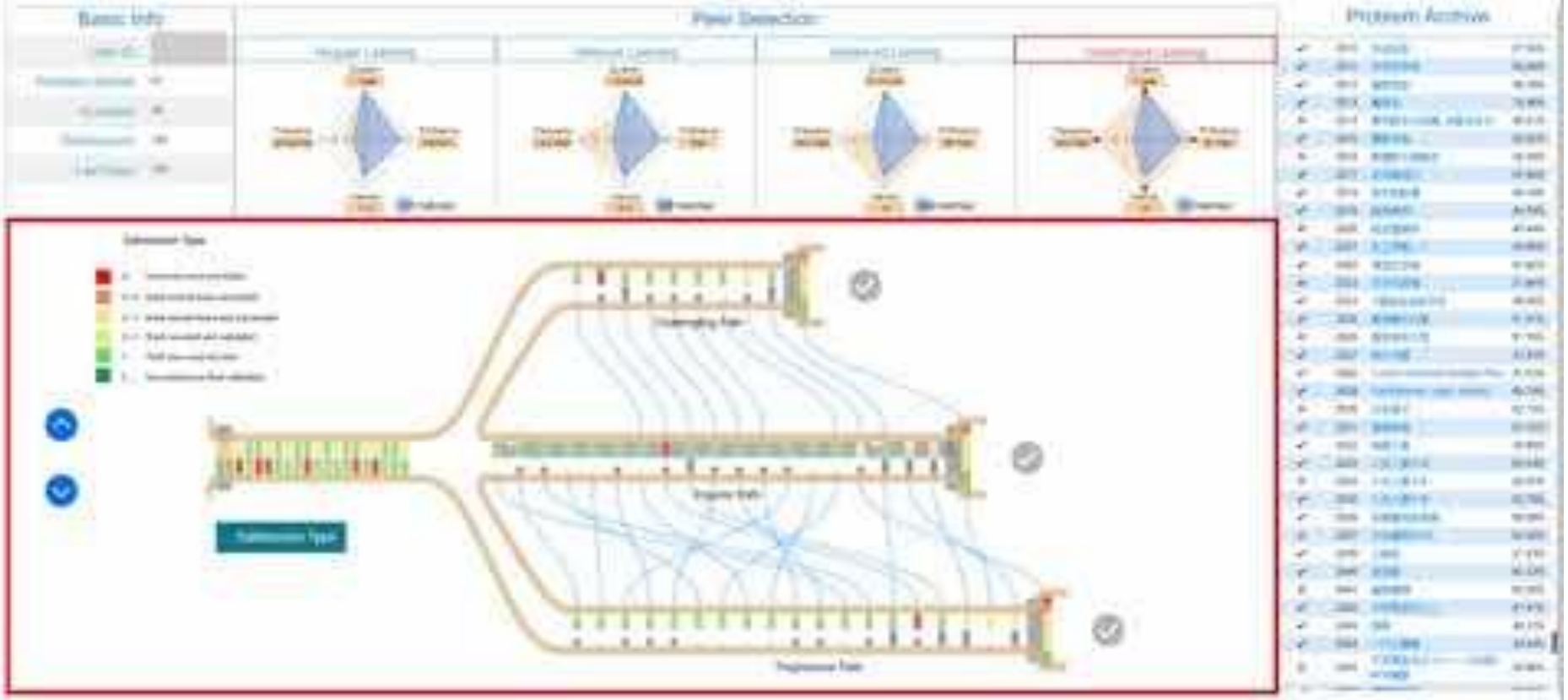


Popular path: $X_1 \rightarrow X_3 \rightarrow X_4$

Challenging path: $X_1 \rightarrow X_4$

Progressive path: $X_3 \rightarrow X_4 \rightarrow X_1$



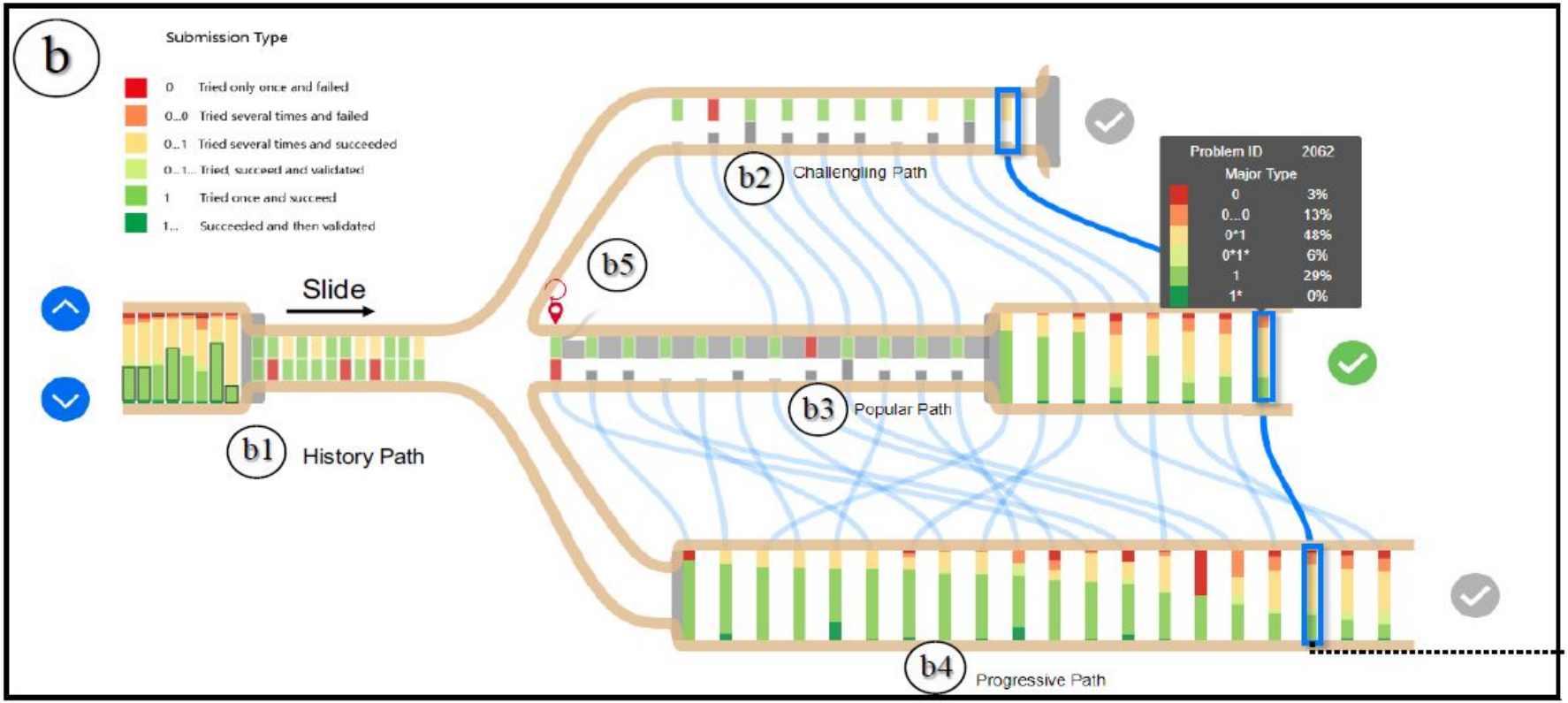

Basic Info
Peer Selection
Problem Archive

Visual Design: Peer Selection View

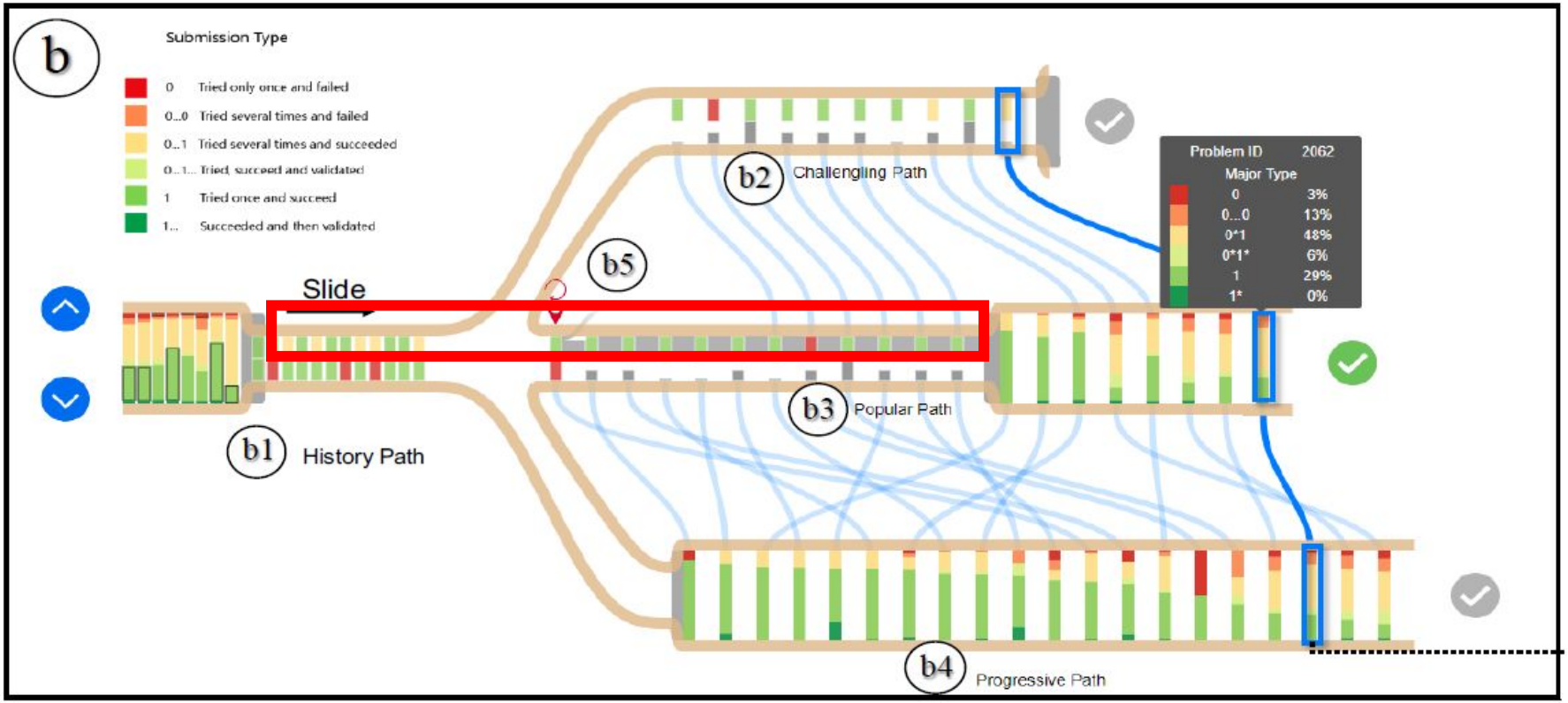
Yellow diamond plot: selected peers **Blue diamond plot: learner himself**

- **Regular Learning**: regularly for a long time and solve 1-2 problems per day.
- **Intensive Learning**: 1-3 months, solve 2-5 questions per day with high proficiency.
- **Advanced Learning**: solve many problems per day in short time with high proficiency.

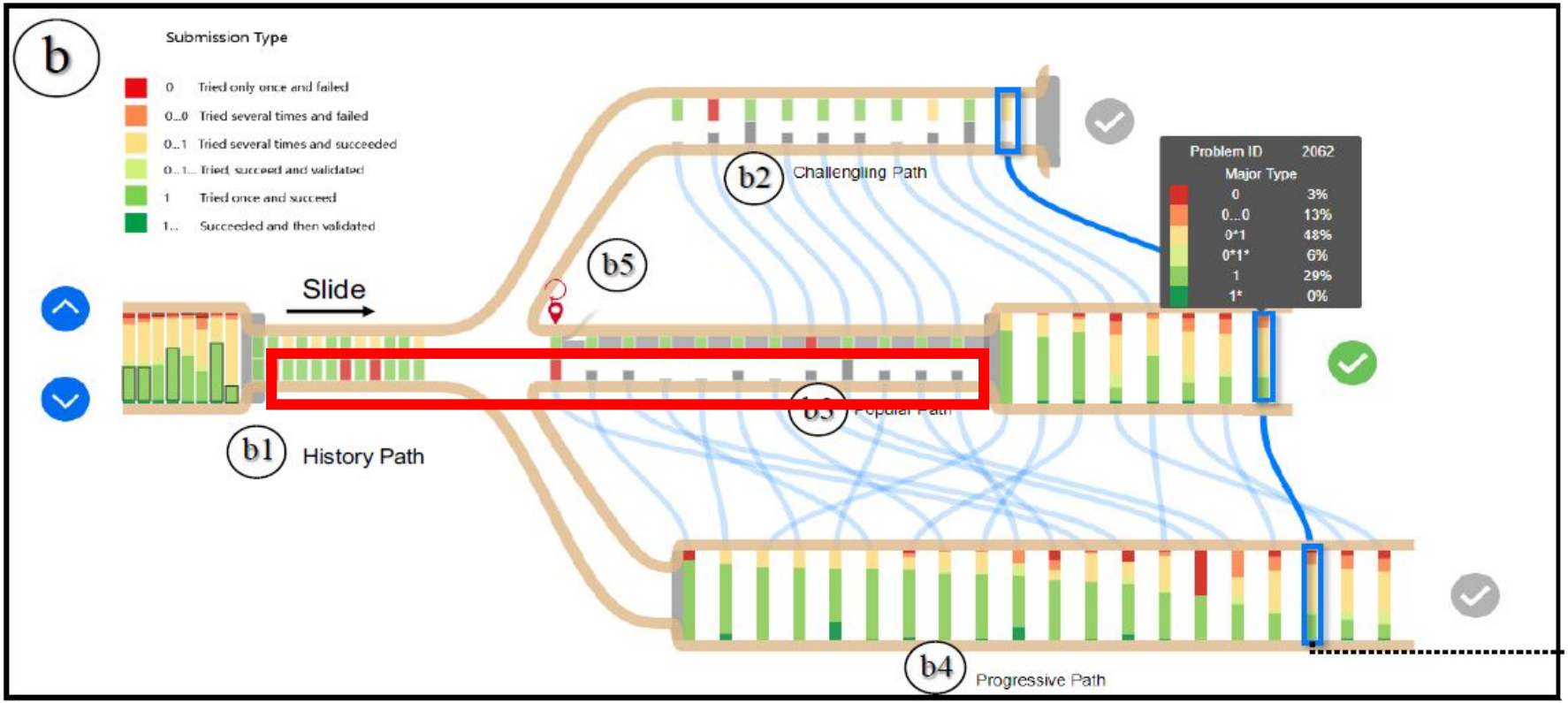
Visual Design: Learning Path View



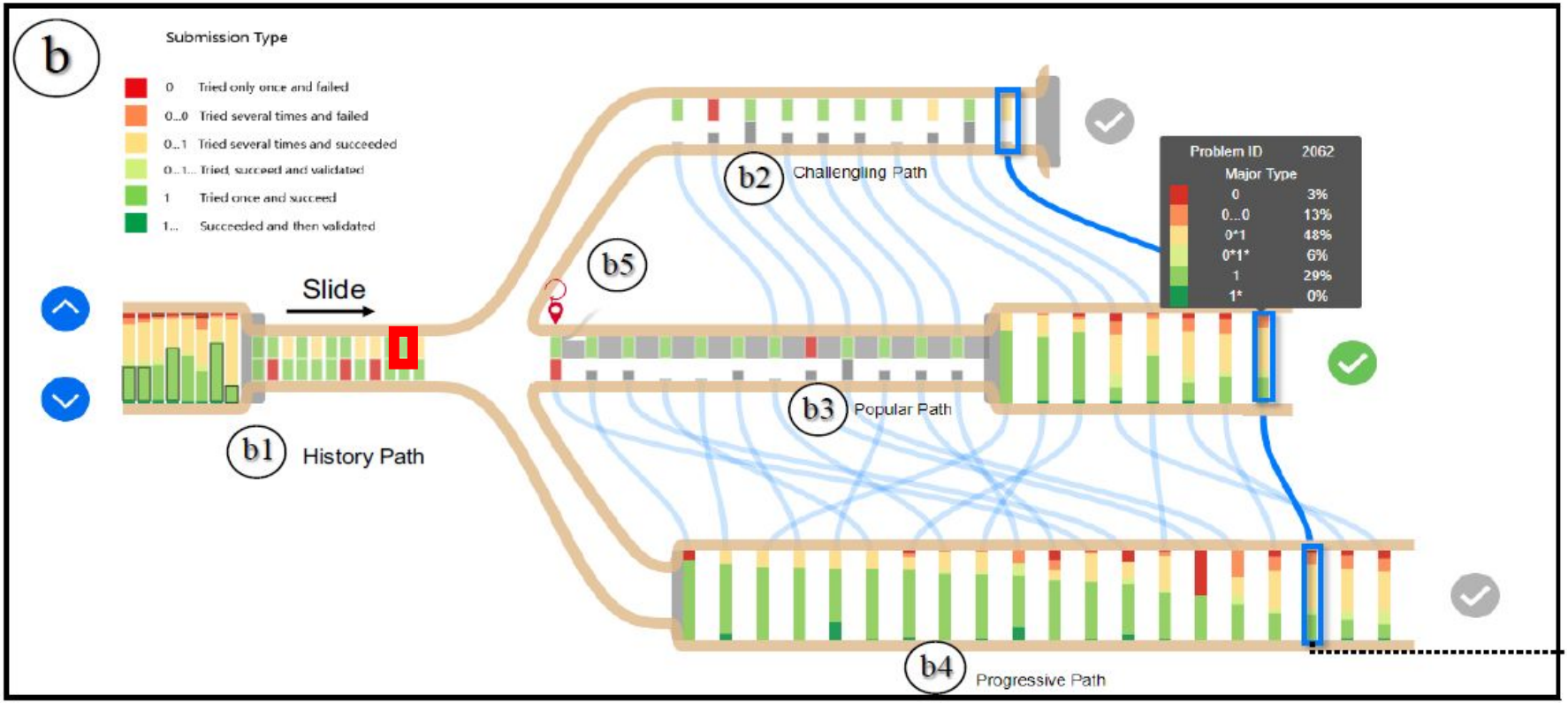
Visual Design: Learning Path View



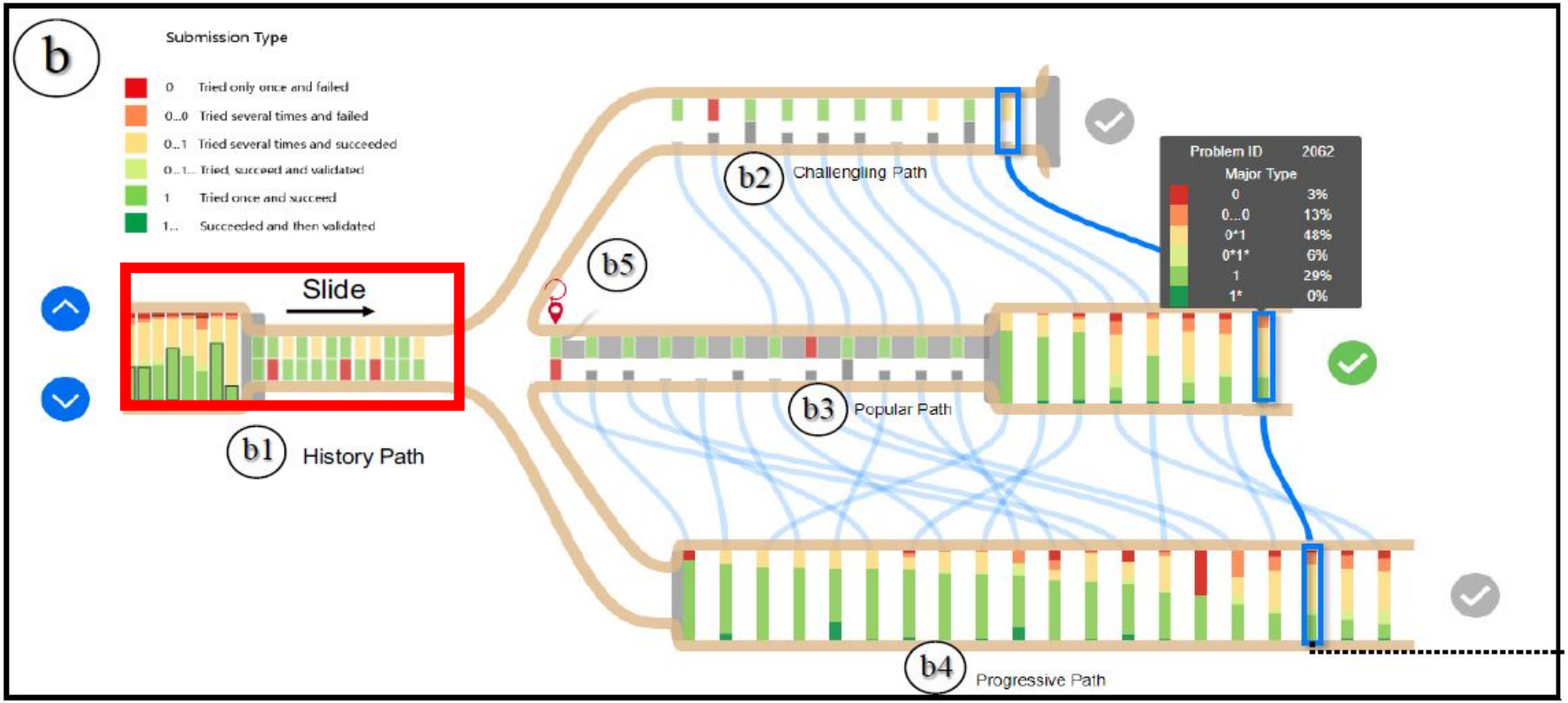
Visual Design: Learning Path View



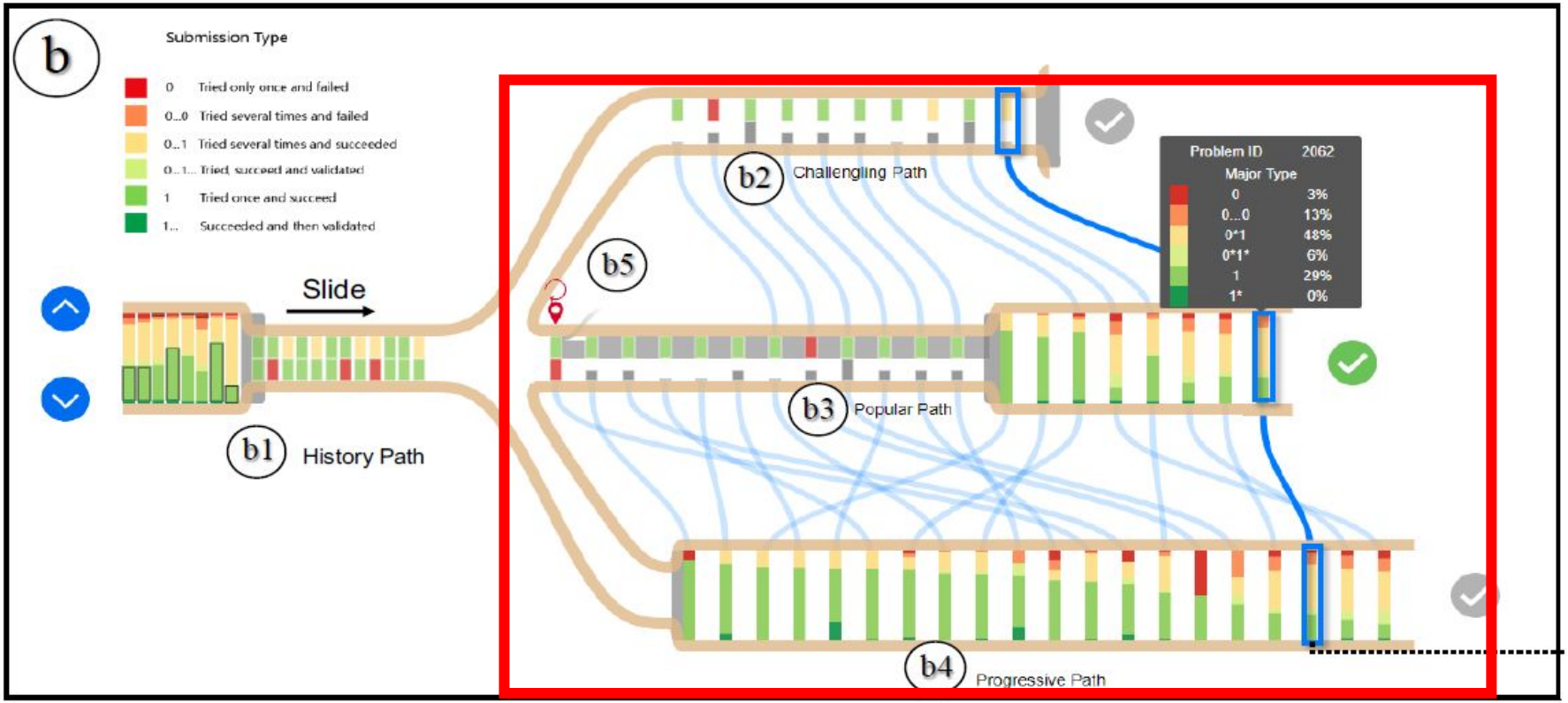
Visual Design: Learning Path View



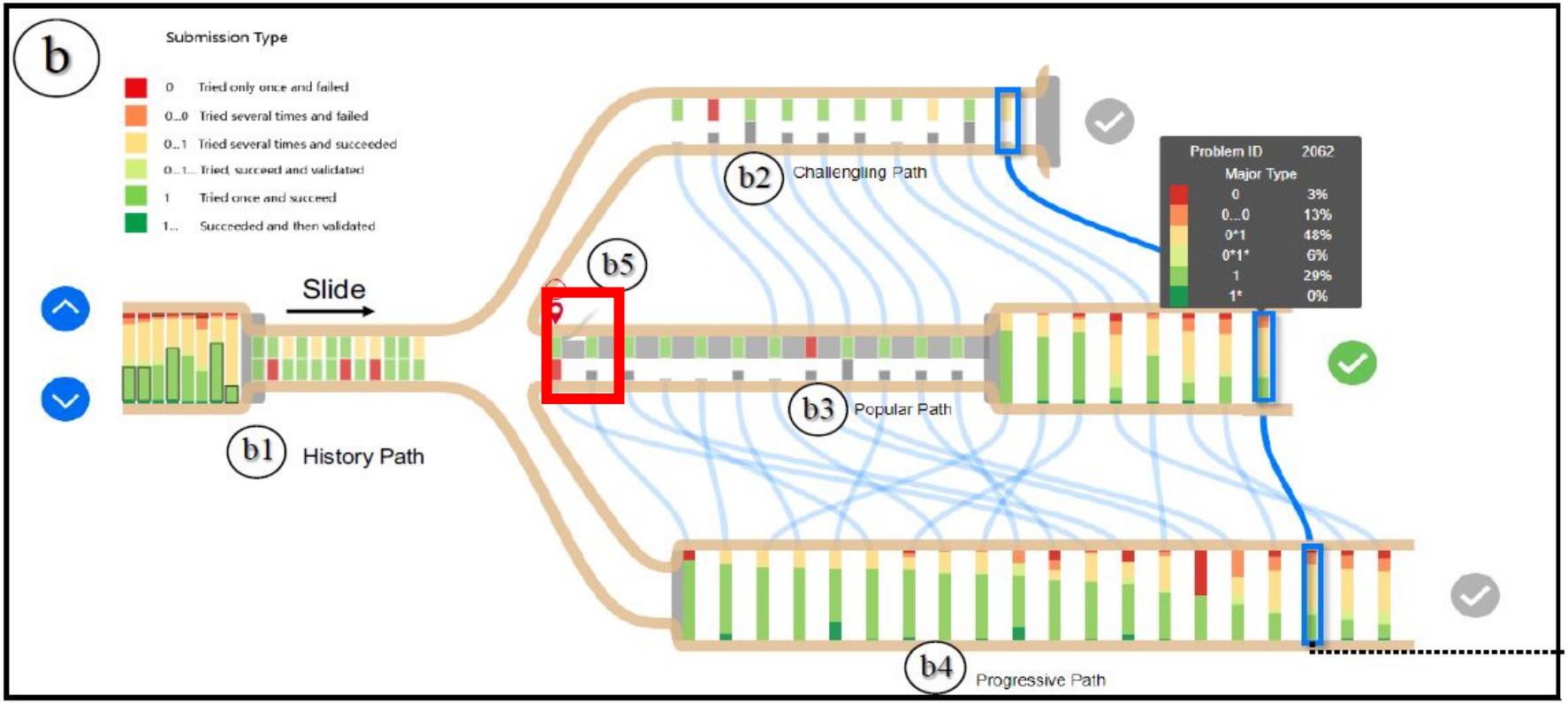
Visual Design: Learning Path View



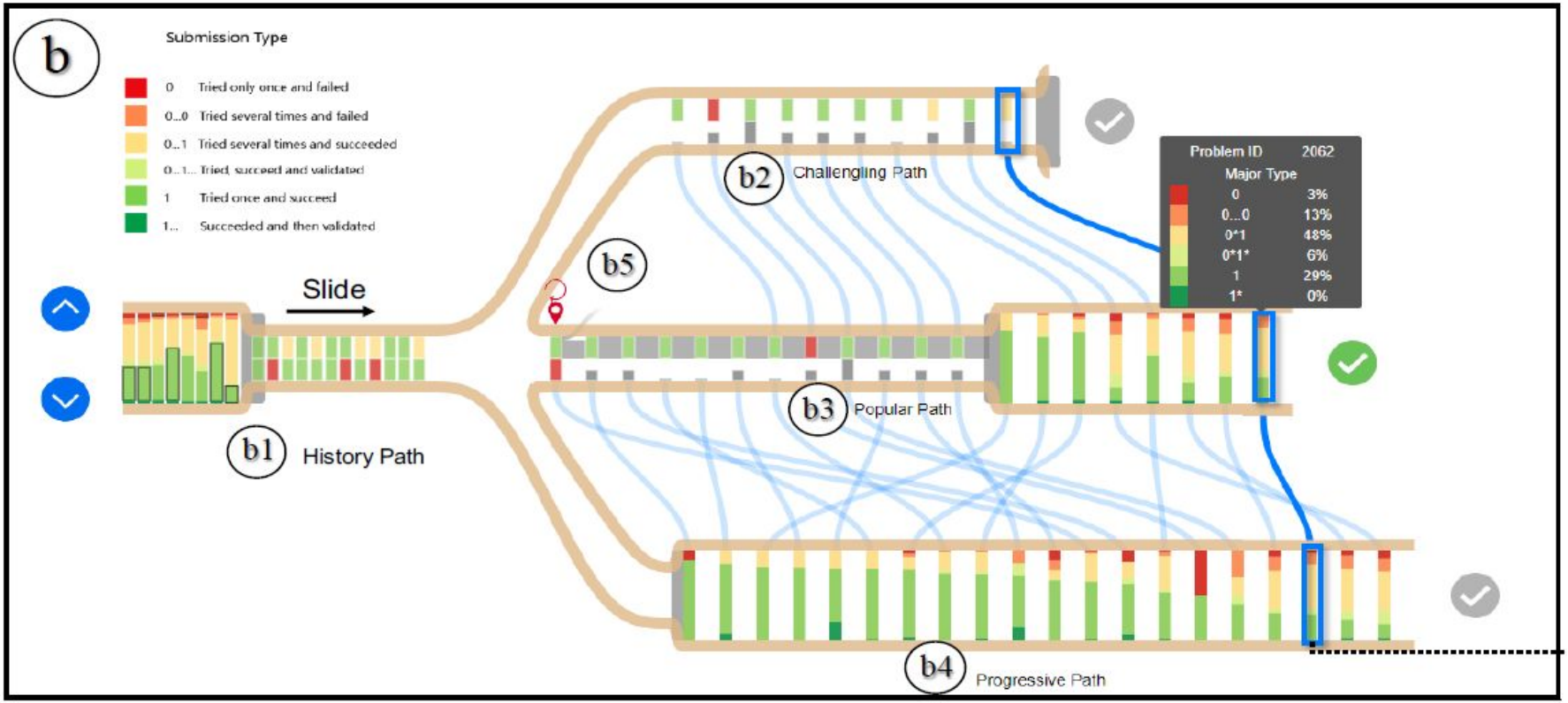
Visual Design: Learning Path View



Visual Design: Learning Path View



Visual Design: Learning Path View



Evaluation: Experiment Design

Dataset:
A popular programming question pool
~4.6M submission records
~54K learners
~5K programming questions

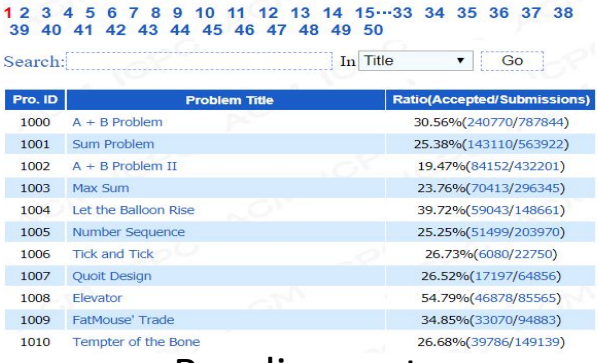
Participants:
18 (7 females, 11 males, age:24±2.85), from a local computer science department

Systems:

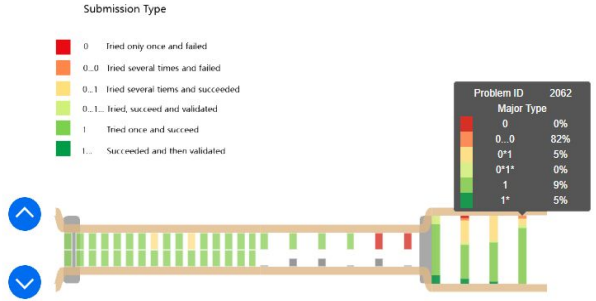
S1. Full PeerLens

S2. Baseline system

S3. Primitive PeerLens



Baseline system



Primitive PeerLens

Evaluation: Experiment Design

Learning scenarios:

L1. Basic programming practice

L2. Coding qualification test for IT company interviews

L3. International Programming Contest

Within-subject:
Counter balance the three learning scenarios and three systems

Tasks:

1. Determine the starting question under a specific learning scenario
2. Find the next question to solve given an existing historical learning path

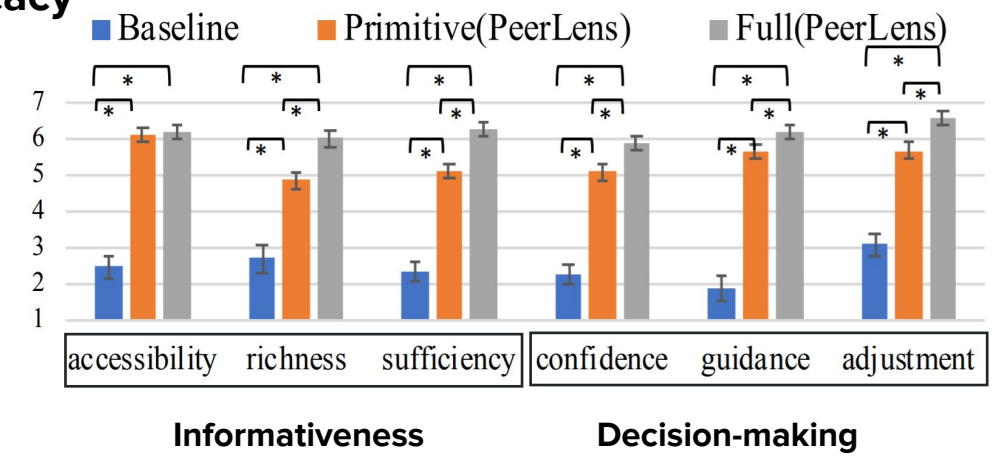
Evaluation: Questionnaires

Informativeness	Q1	The information needed to plan a learning path is easy to access.
	Q2	The information needed to plan a learning path is rich.
	Q3	The information is sufficient to plan a learning path.
Decision making	Q4	The system was helpful for me to find a proper learning path for a specific learning scenario.
	Q5	I am confident that I find a suitable learning path for the learning scenario.
	Q6	The system helps make adjustment according to previous performance.
Visual design	Q7	The learning path design is intuitive.
	Q8	The learning path design helps me understand the suggested path.
System Usability	Q9	It was easy to learn the system.
	Q10	It was easy to use the system.
	Q11	I would like to recommend this system to others.

Results

Informativeness and decision-making efficacy

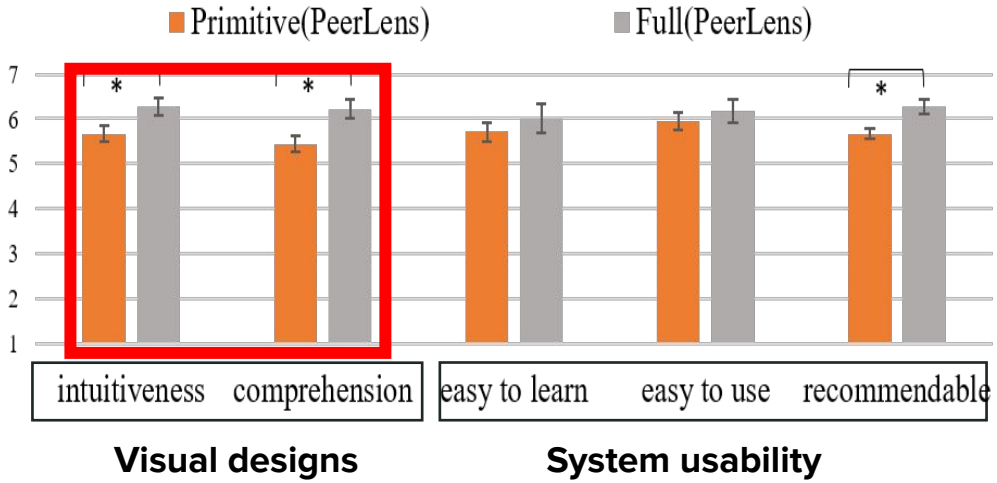
- Primitive and Full PeerLens > Baseline
- Information richness & sufficiency:
  Full PeerLens > Primitive
- Information accessibility:
  No significant differences between Full and Primitive
- Decision-making metrics:
  Full PeerLens > Primitive



Results

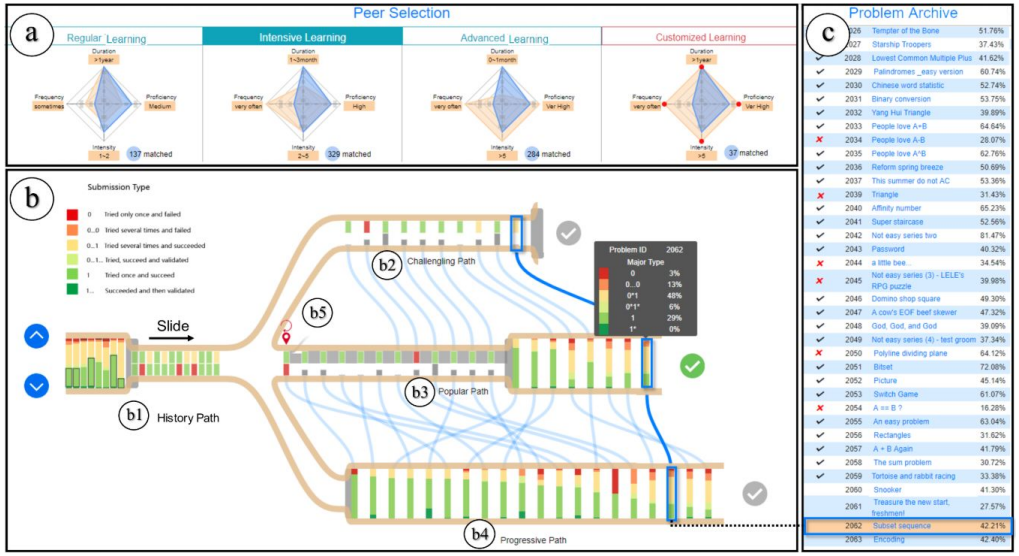
Visual designs and system usability

- Intuitiveness & comprehension:
  Full PeerLens > Primitive
- Easy to learn & use:
  No significant difference
  between Full and Primitive
- Recommendation:
  Full PeerLens > Primitive



Conclusion

- A novel visual analytics system for comparing event sequence data and providing explanation for recomemendation
- A novel zipper-like visualization for showing information step by step to reduce cognitive load
- A within-subject user experiment to show the system usefulness and usability



My works

Learners (learning loop)

Educators (design loop)

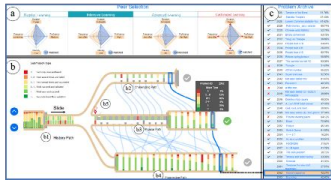
Customizing
learning goals and personalize activities

Reflecting
self-regulations on learning behaviors

Analyzing
learners behaviors and improve learning design

Data

- Large heterogeneous data
- Limited expertise and time in data analysis
- Not enough guidance and explanations
- No guarantee of data quality

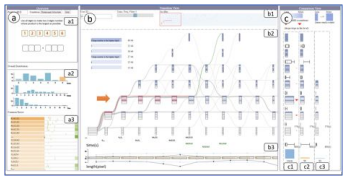


Peerlens (CHI 2019)

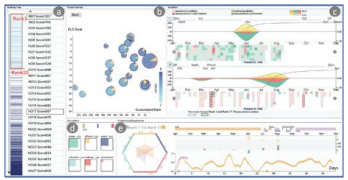
- Large heterogeneous data
- Limited expertise and time in data analysis
- Lack of motivation, consistent mental model, and actionable plans

RLens (L@S 2022)

- Large heterogeneous data
- Limited expertise and time in data analysis
- No predefined model



QLens (TVCG 2021)



SeqDynamics (EuroVIS 2020)

RLens: A Computer-aided Visualization System for Reflecting Language Learning Progress Under Distributed Tutorship

Meng Xia, Yankun Zhao*, Jihyeong Hong*, Mehmet Hamza Erol*, Taewook Kim, Juho Kim

L@S 2022

Background

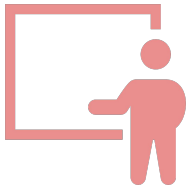
- **Gig economy** gains popularity
  - Temporary, flexible jobs are commonplace for efficient resource allocation
- New modes of teaching and learning spring up

Background

In particular, online language tutoring platforms (e.g., Cambly,) are becoming increasingly popular.

Background

These online language tutoring platforms:

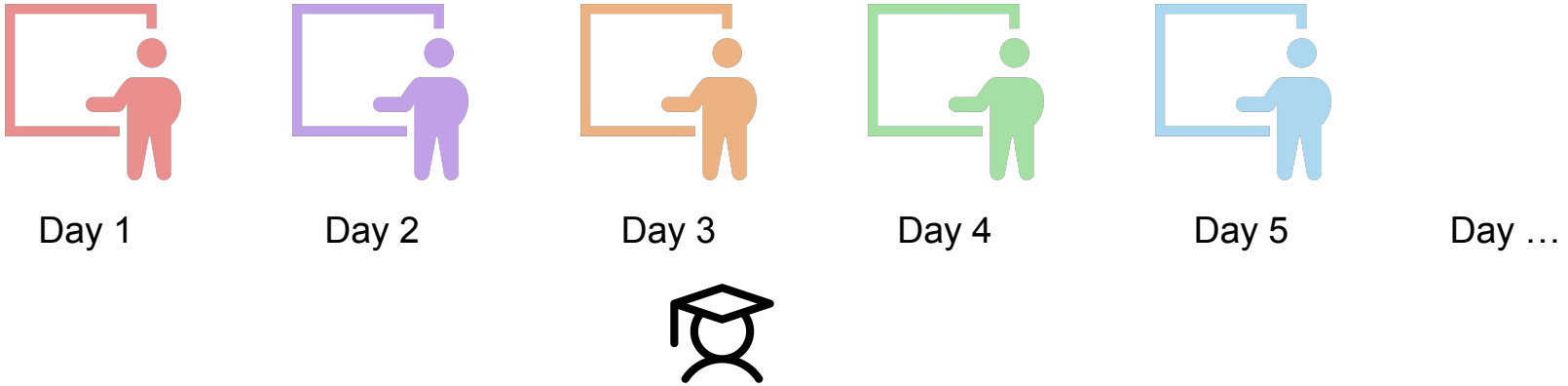


Provide temporary jobs for native speakers to work as part-time tutors

Enable language learners to have 1-1 speaking sessions with native speakers anytime and anywhere

Distributed Tutorship

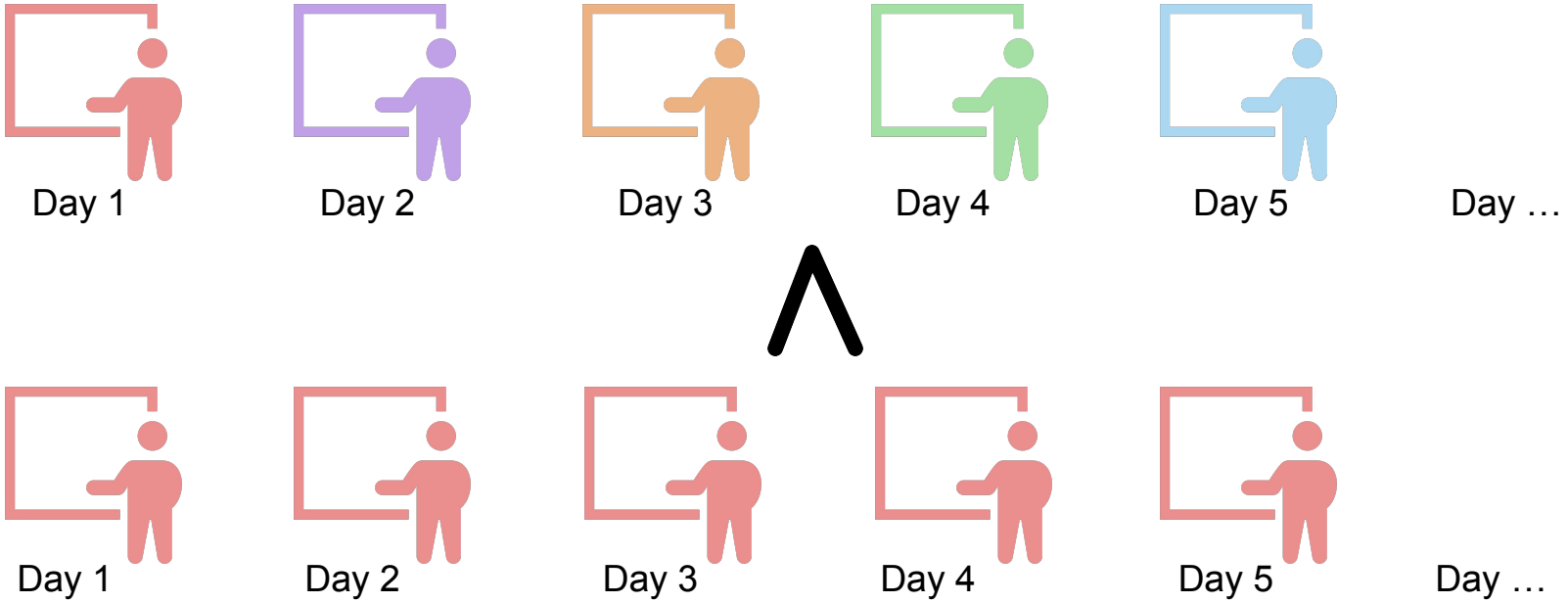
Our previous work [1] analyzed 15,959 learners' data on one of these platforms and identified that **learners actively distribute their learning time with different tutors during the learning process,** which was defined **as distributed tutorship**.



[1] Xia, Meng, et al. "Understanding Distributed Tutorship in Online Language Tutoring." *LAK22: 12th International Learning Analytics and Knowledge Conference*. 2022.

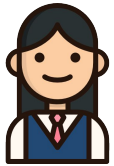
Distributed Tutorship

There is suggestive evidence that **more distributed tutorship might introduce lower learning improvement** [1].



Needs-finding Interviews

An online English tutoring platform. On Ringle, learners can choose tutors and class time for 1:1 online speaking sessions.

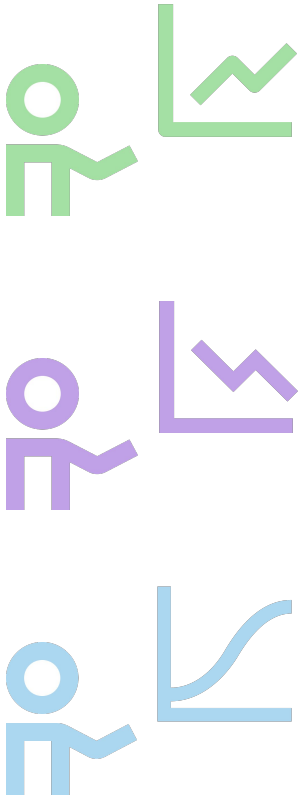


16 learners, who have learnt from more than one tutor.

For example, how do you calibrate your progress when you have taken multiple sessions with different tutors? Have you encountered any difficulties?

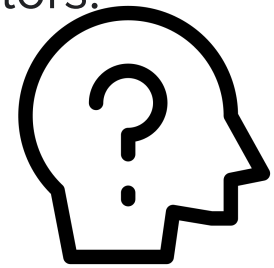
Challenge #1: Grading Inconsistency

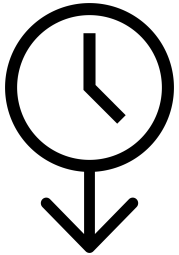
Learners have a hard time knowing their improvement through checking the scores of different tutors, since each tutor might have different grading standards.

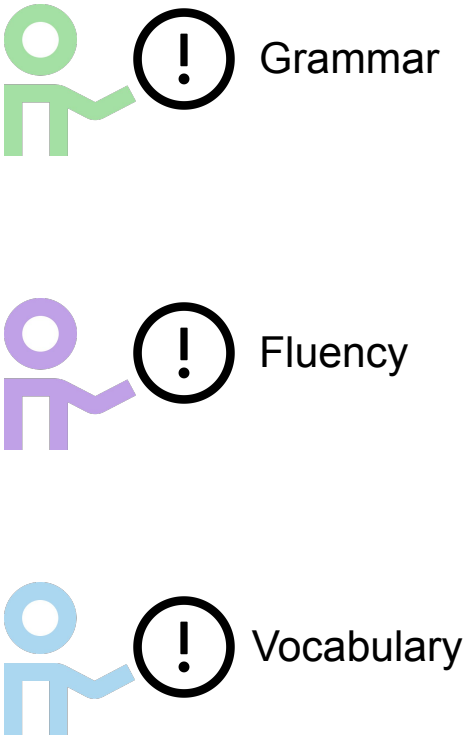


Challenge #2: Feedback Discontinuity

Learners are unaware of their common language issues (e.g. tense errors) and they are not sure whether they have corrected the issues or not, since previous corrections are not tracked by different tutors.

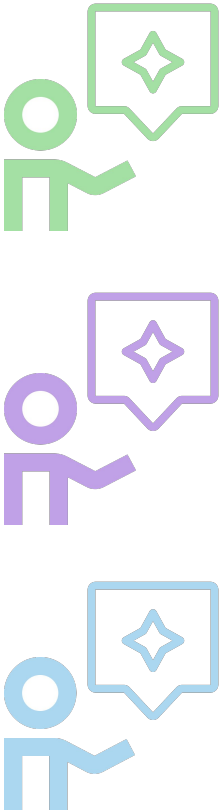






Challenge #3: Unorganized Feedback

Learners are uncertain about what to do next, since suggestions given by different tutors are from diverse perspectives.



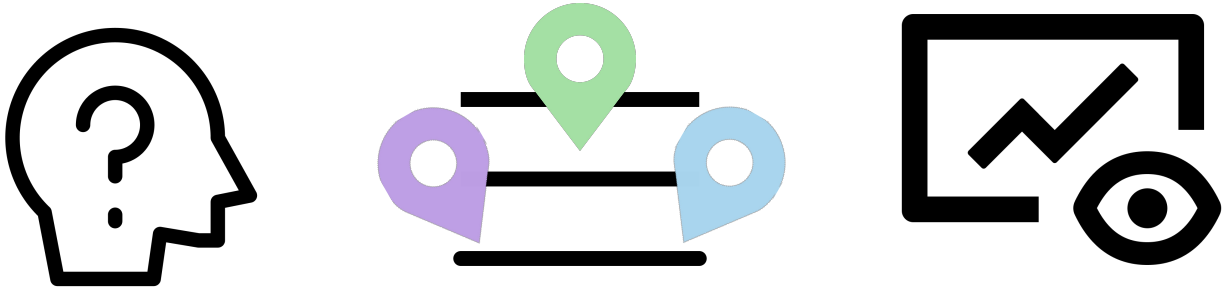
Challenge 4#: Lack of Context for Feedback Understanding

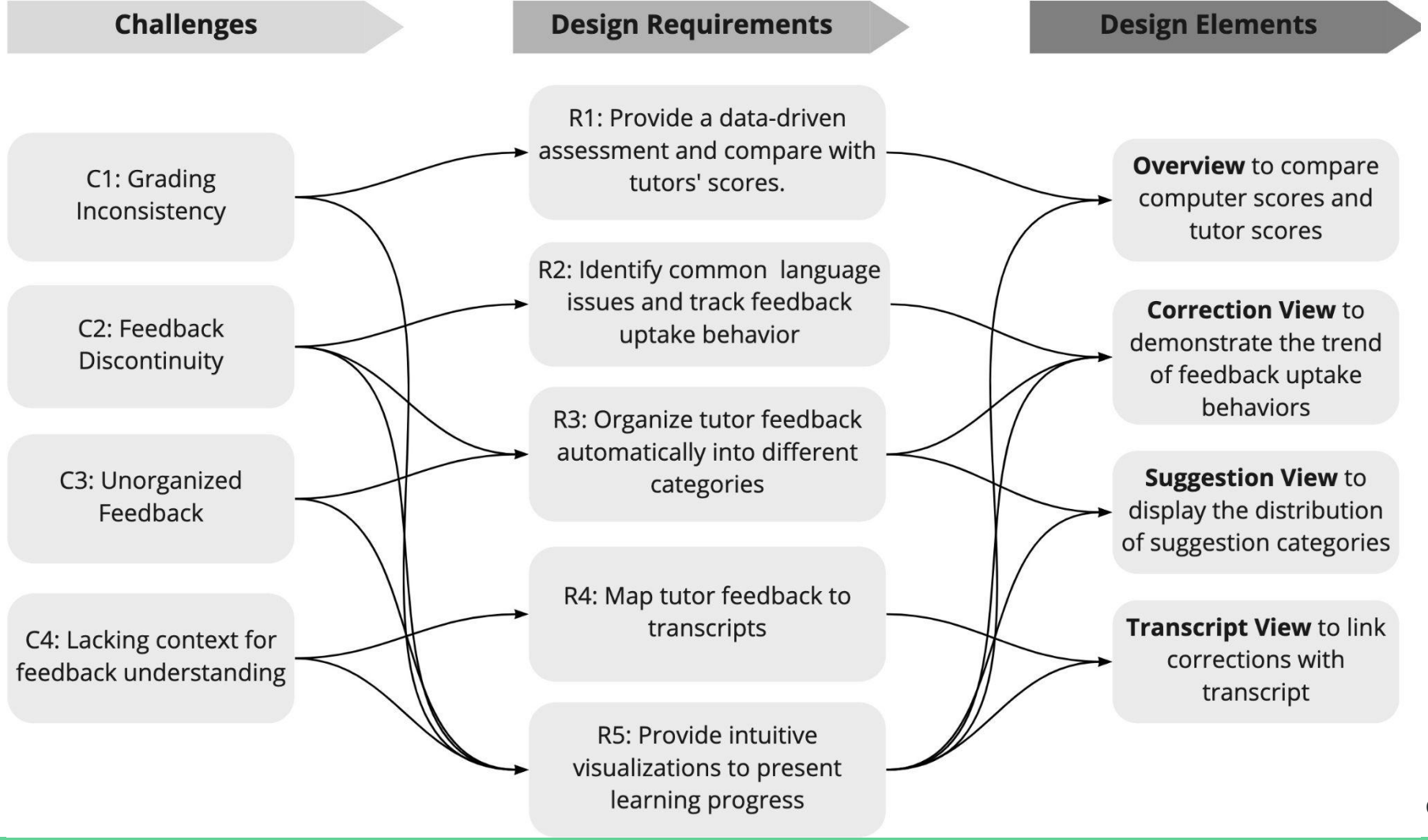
It would take a long time to find the corresponding place in the audio recording where the feedback was given.

We want to design **intuitive** learning dashborad

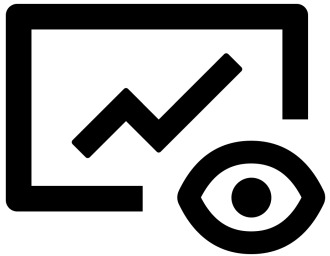
Learners are not data scientists. Exhaustive visual analysis may not be their best option.

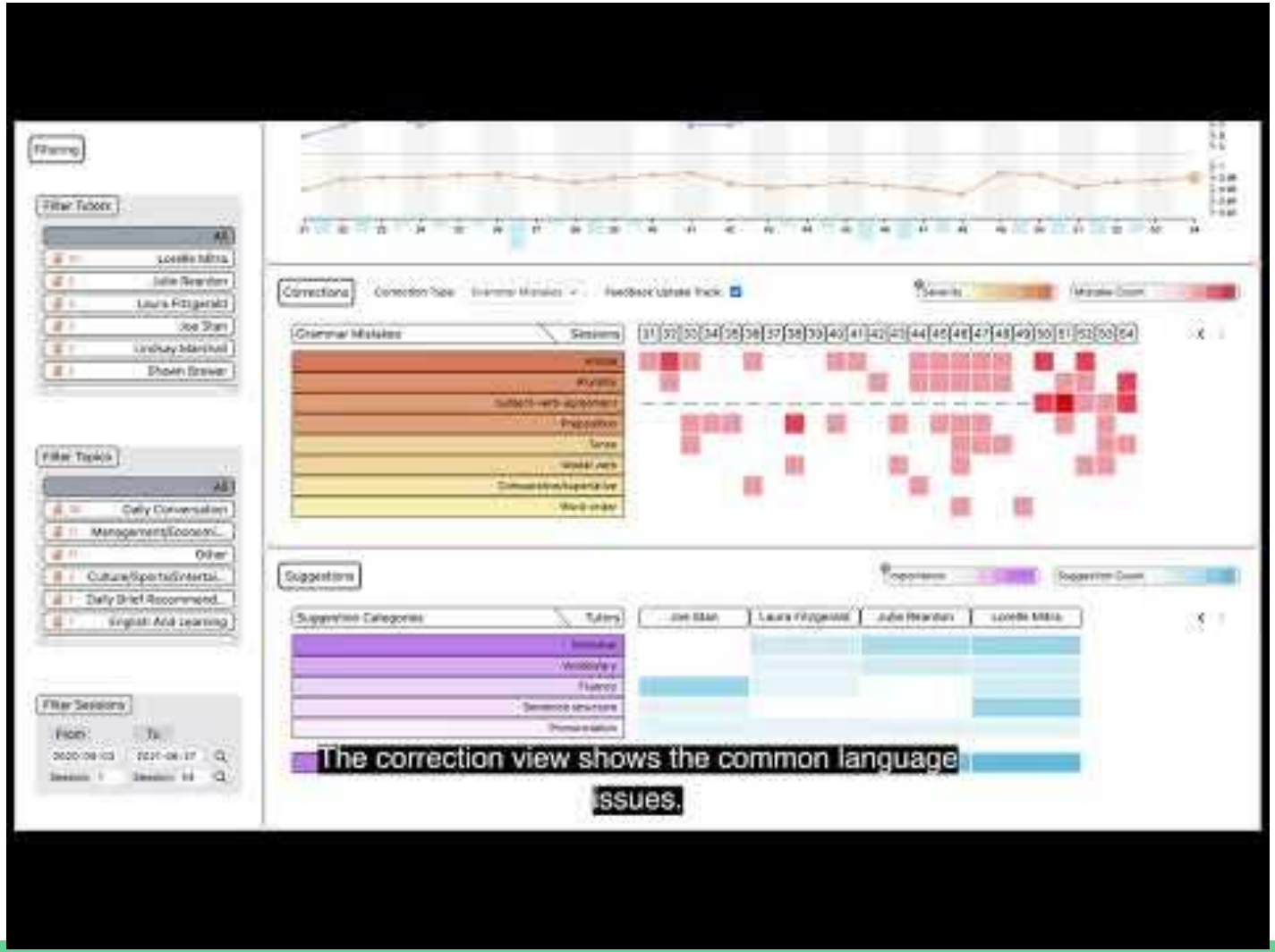
We need to build an **intuitive** visual analysis system to reduce the cognitive load

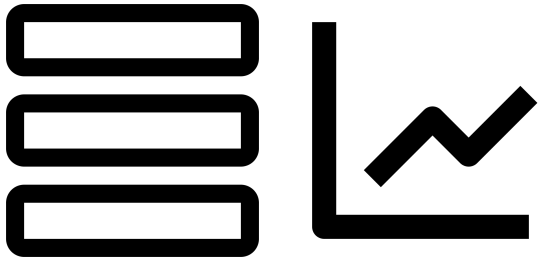



Challenges
Design Requirements
Design Elements
C1: Grading Inconsistency
C2: Feedback Discontinuity
C3: Unorganized Feedback
C4: Lacking context for feedback understanding
R1: Provide a data-driven assessment and compare with tutors' scores.
R2: Identify common language issues and track feedback uptake behavior
R3: Organize tutor feedback automatically into different categories
R4: Map tutor feedback to transcripts
R5: Provide intuitive visualizations to present learning progress
Overview to compare computer scores and tutor scores
Correction View to demonstrate the trend of feedback uptake behaviors
Suggestion View to display the distribution of suggestion categories
Transcript View to link corrections with transcript

NLP +

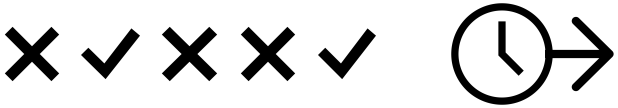



Filtering
Filter Tutors
All
Lorelle Mitra
Julie Reardon
Laura Fitzgerald
Joe Stan
Lindsay Marshall
Shawn Brewer
Filter Topics
All
Daily Conversation
Management/Economi...
Other
Culture/Sports/Entertai...
Daily Brief Recommend...
English And Learning
Filter Sessions
From
To
Corrections
Grammar Mistakes
Sessions
Suggestions
Suggestion Categories
Tutors
Joe Stan
Laura Fitzgerald
Julie Reardon
Lorelle Mitra
The correction view shows the common language
issues.

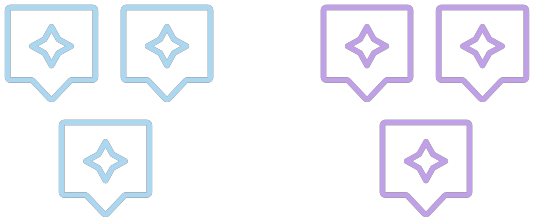


Data Driven Evaluation

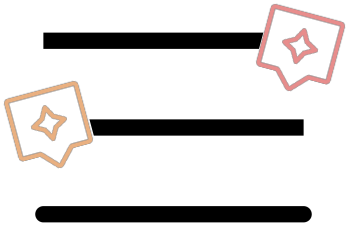
Grammar



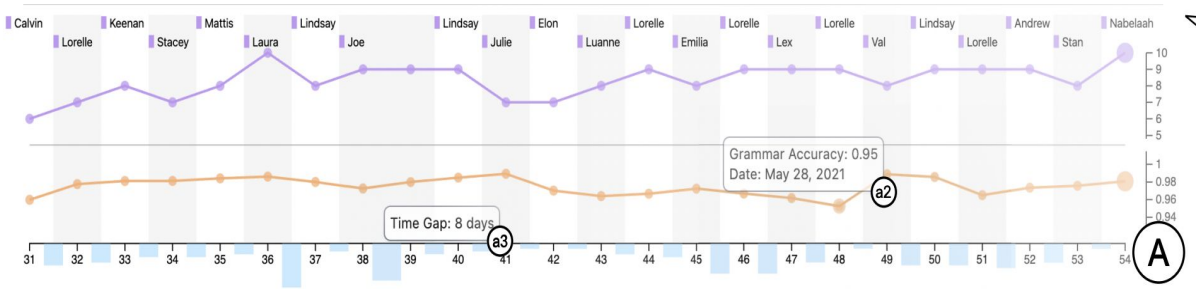
Identify and Track Common Language Issues



Feedback Categorization



Feedback-Context Mapping



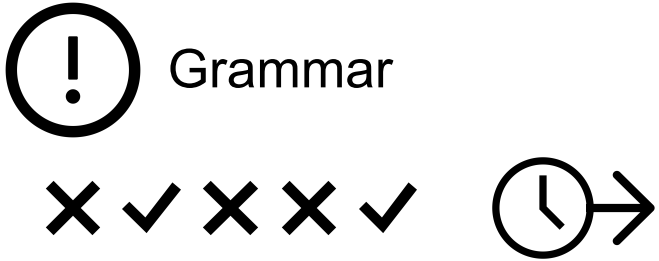
Data Driven Evaluation

Vocabulary Complexity: measure of textual lexical diversity (MTLD), the average length of sequential words a speaker can produce that keep the type-token ratio (TTR) higher than 0.72.

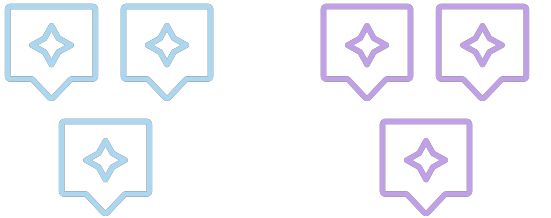
Grammar Accuracy: ratio of error-free C-Units to the total number of C-Units, where C-Unit is defined as the minimal communication unit (e.g., ``Yes.")

Fluency: Mean Length of Run (MLR), the average number of syllables per utterance without any pause, where the threshold for pause identification is set to 250 ms.

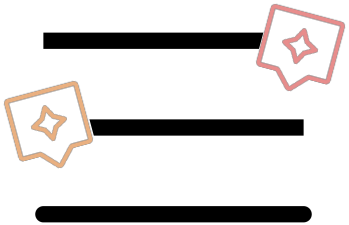
Data Driven Evaluation



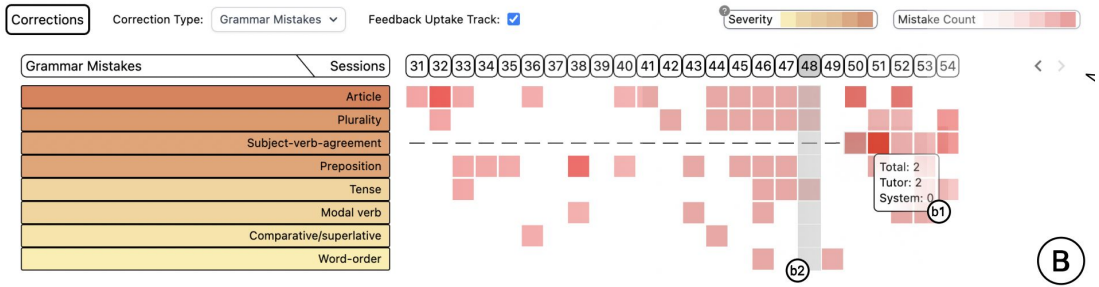
Identify and Track Common Language Issues

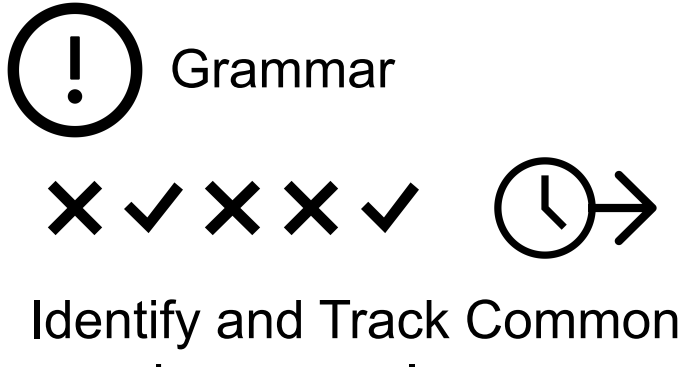


Feedback Categorization



Feedback-Context Mapping





Corrective feedback: (e.g., two apple -> two apples, think positively -> optimistic)

Feedback uptake behaviors

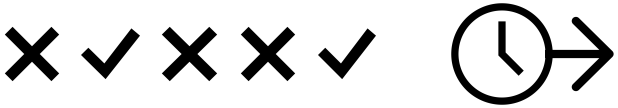
Grammar: whether still have the grammar issue mentioned by previous tutors

Vocabulary: for each session, check using **masked language modeling**

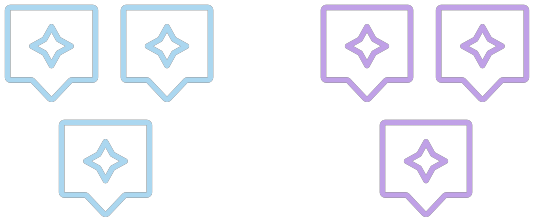
- whether the suggested expressions have been used correctly
- whether the original expressions were still used incorrectly

Data Driven Evaluation

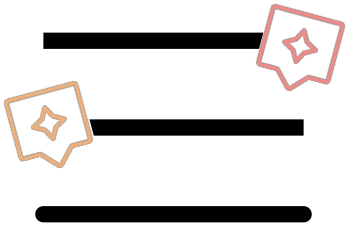
Grammar



Identify and Track Common Language Issues

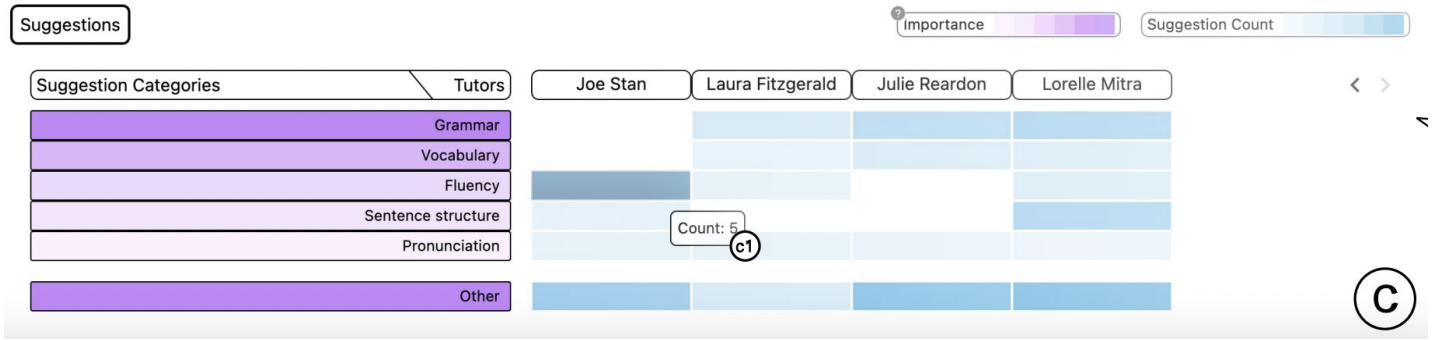


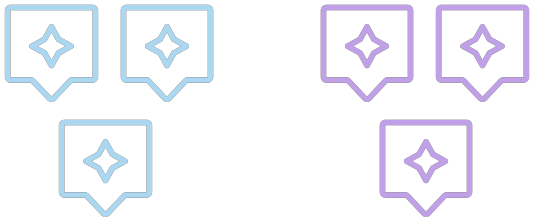
Feedback Categorization



Feedback-Context Mapping

(1) **Manually selected six categories** with the help of three tutors
(2) **Sentence classification** for each feedback sentence using **natural language inference techniques**

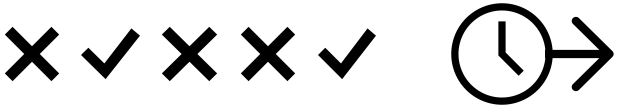




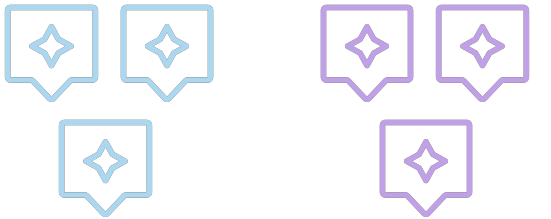
Feedback Categorization

Data Driven Evaluation

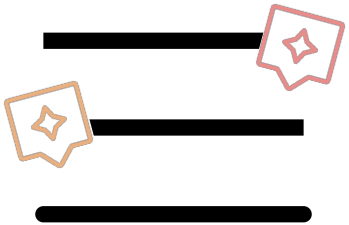
Grammar



Identify and Track Common Language Issues

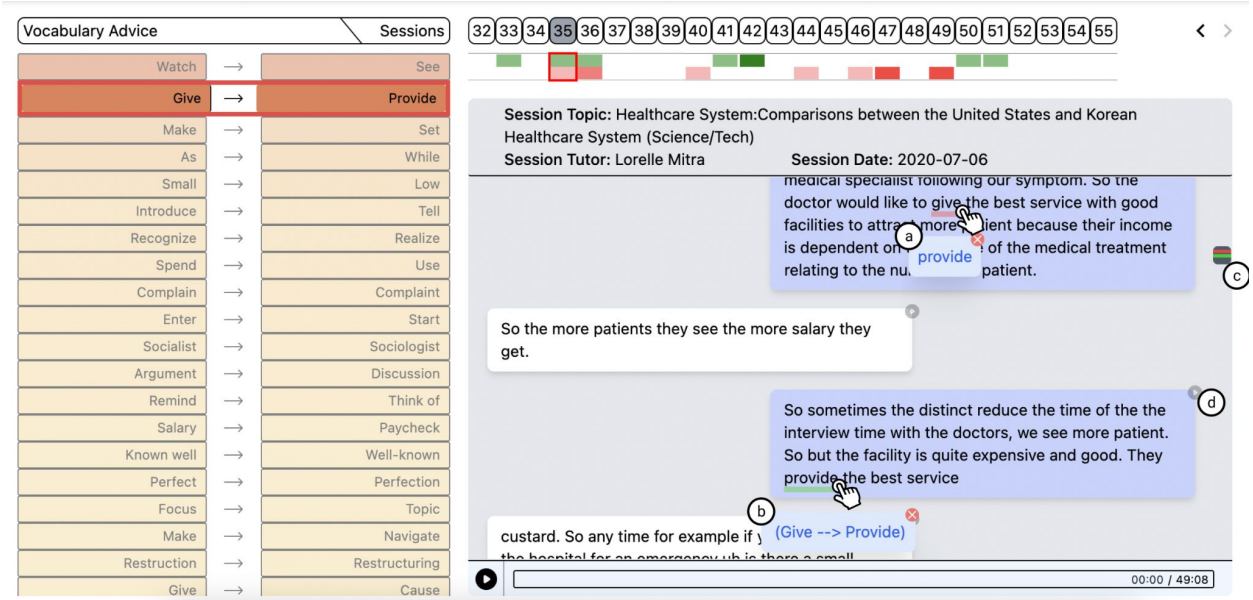


Feedback Categorization



Feedback-Context Mapping

Map tutors' feedback to the transcripts based on the sentence similarity



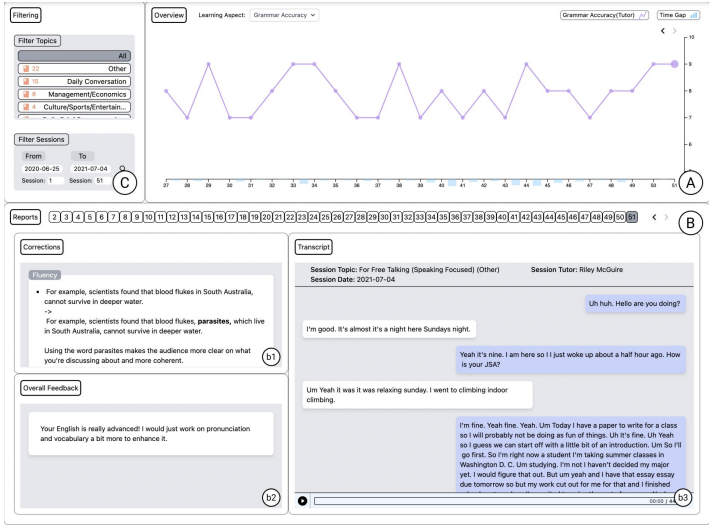
Feedback-Context Mapping

User Study

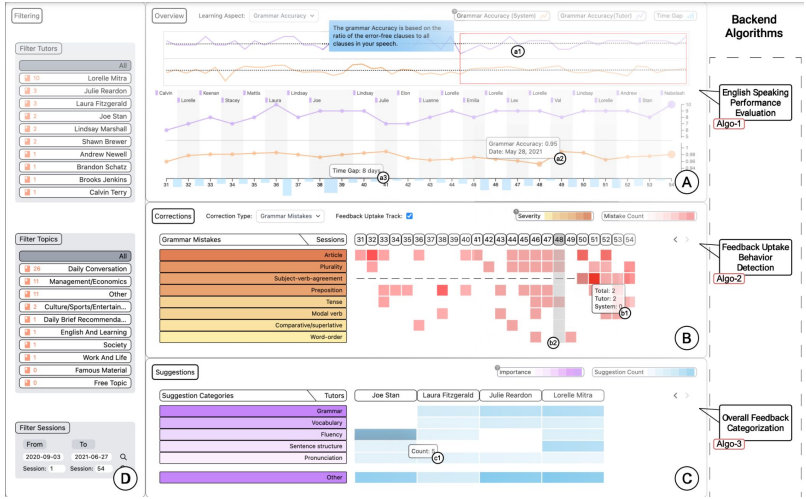
A between-subjects study on a Baseline system and RLens.

40 learners from Ringle, who have learnt from more than one tutor and 25 sessions.

User Study



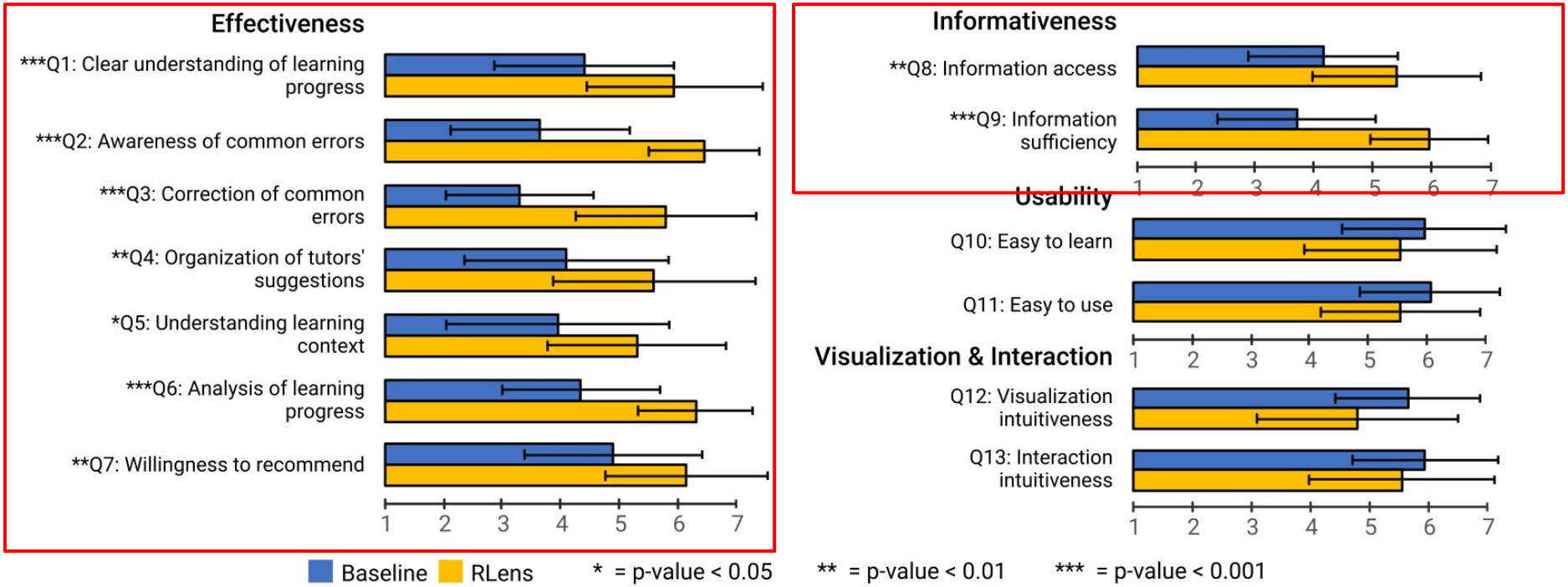
Baseline: 20 learners

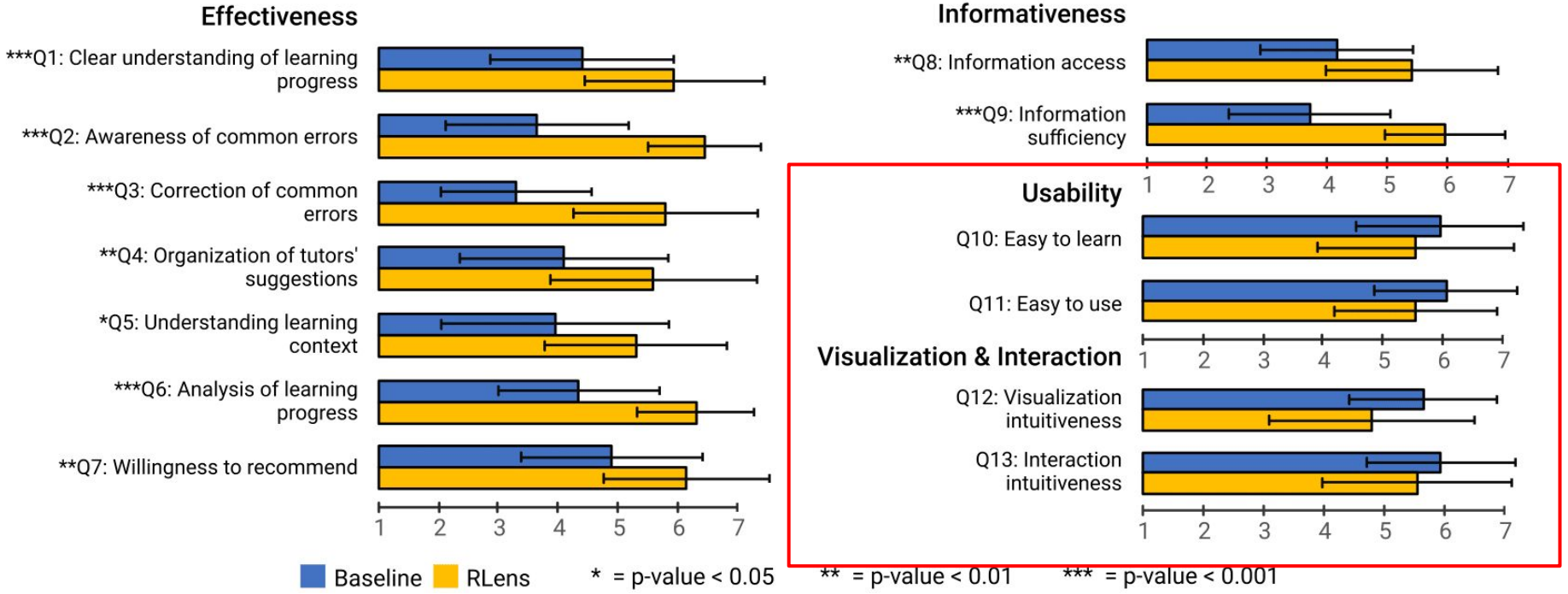


RLens: 20 learners

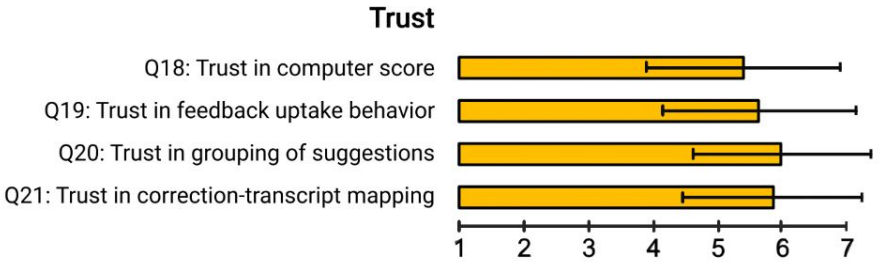
User Study

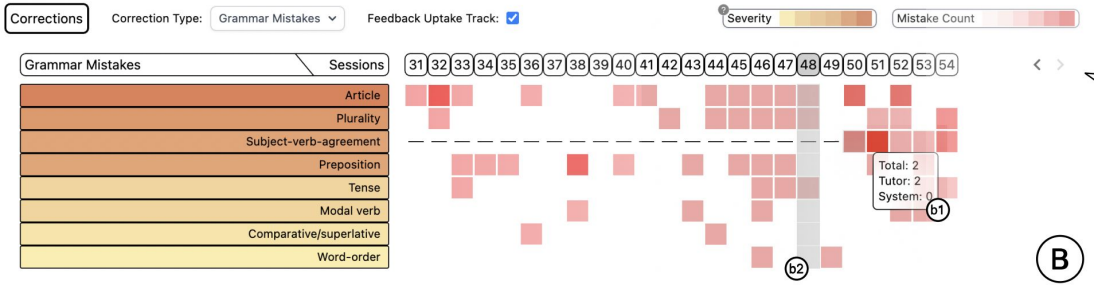
T1: Please describe your overall learning progress.
T2: Please identify your common language issues in the learning process.
T3: Please describe whether you have corrected your common language issues in the learning process.
T4: Please describe the common aspects in tutors' overall feedback.
T5: Please describe how you check the transcript using the system for learning.
T6: Please describe the reasons for ups and downs in scores showing in Overview.
T7: Please describe how you will use this system in learning reflection if it is deployed.


Effectiveness
***Q1: Clear understanding of learning progress
***Q2: Awareness of common errors
***Q3: Correction of common errors
**Q4: Organization of tutors' suggestions
*Q5: Understanding learning context
***Q6: Analysis of learning progress
**Q7: Willingness to recommend
1 2 3 4 5 6 7
Informativeness
**Q8: Information access
***Q9: Information sufficiency
1 2 3 4 5 6 7
Usability
Q10: Easy to learn
Q11: Easy to use
1 2 3 4 5 6 7
Visualization & Interaction
Q12: Visualization intuitiveness
Q13: Interaction intuitiveness
1 2 3 4 5 6 7
Baseline
RLens
* = p-value < 0.05
** = p-value < 0.01
*** = p-value < 0.001


Effectiveness
***Q1: Clear understanding of learning progress
***Q2: Awareness of common errors
***Q3: Correction of common errors
**Q4: Organization of tutors' suggestions
*Q5: Understanding learning context
***Q6: Analysis of learning progress
**Q7: Willingness to recommend
1 2 3 4 5 6 7
Informativeness
**Q8: Information access
***Q9: Information sufficiency
1 2 3 4 5 6 7
Usability
Q10: Easy to learn
Q11: Easy to use
1 2 3 4 5 6 7
Visualization & Interaction
Q12: Visualization intuitiveness
Q13: Interaction intuitiveness
1 2 3 4 5 6 7
Baseline
RLens
* = p-value < 0.05
** = p-value < 0.01
*** = p-value < 0.001


Visualization Intuitiveness
Q14: Overview intuitiveness
Q15: Correction View intuitiveness
Q16: Suggestion View intuitiveness
Q17: Transcript View intuitiveness
1
2
3
4
5
6
7


Trust
Q18: Trust in computer score
Q19: Trust in feedback uptake behavior
Q20: Trust in grouping of suggestions
Q21: Trust in correction-transcript mapping
1
2
3
4
5
6
7

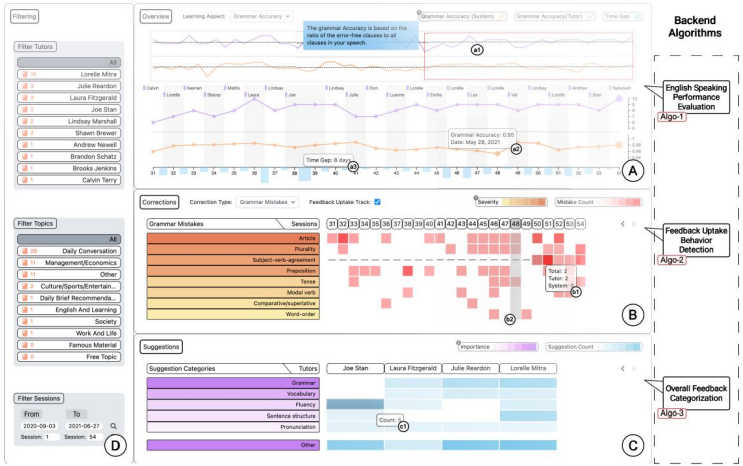

Corrections
Correction Type:
Grammar Mistakes
Feedback Uptake Track:
Severity
Mistake Count
Grammar Mistakes
Sessions
31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54
Article
Plurality
Subject-verb-agreement
Preposition
Tense
Modal verb
Comparative/superlative
Word-order
Total: 2
Tutor: 2
System: 0
b1
b2
B

Contributions

- **A computer-aided visualization system** for analyzing audio/text learning data to facilitate learners' reflection on the learning process under **distributed tutorship**

- **A user study** showing the effectiveness of reflecting learning progress with RLens

- **A set of design considerations** for computer-aided learning systems under distributed tutorship, e.g., surfacing actionble information



My works

Learners (learning loop)

Educators (design loop)

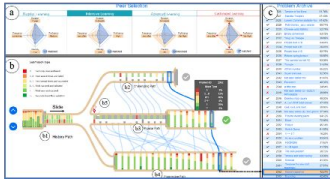
Customizing
learning goals and personalize activities

Reflecting
self-regulations on learning behaviors

Analyzing
learners behaviors and improve learning design

Data

- Large heterogeneous data
- Limited expertise and time in data analysis
- Not enough guidance and explanations
- No guarantee of data quality

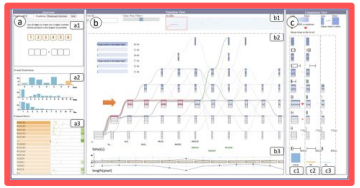


Peerlens (CHI 2019)

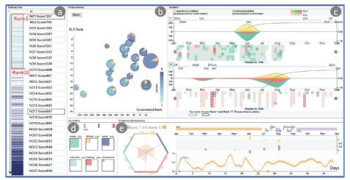
- Large heterogeneous data
- Limited expertise and time in data analysis
- Lack of motivation, consistent mental model, and actionable plans

RLens (L@S 2022)

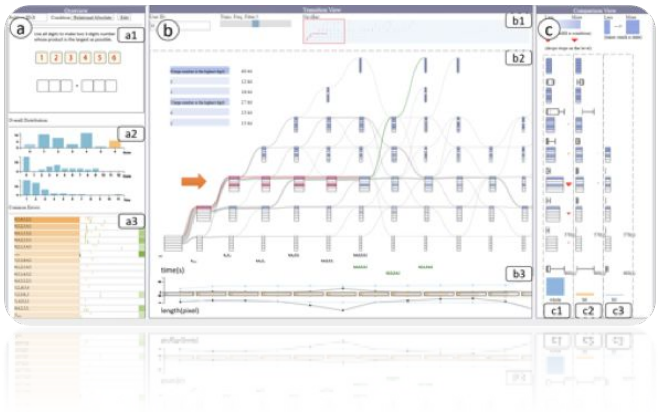
- Large heterogeneous data
- Limited expertise and time in data analysis
- No predefined model



QLens
(TVCG 2021)



SeqDynamics
(EuroVIS 2020)

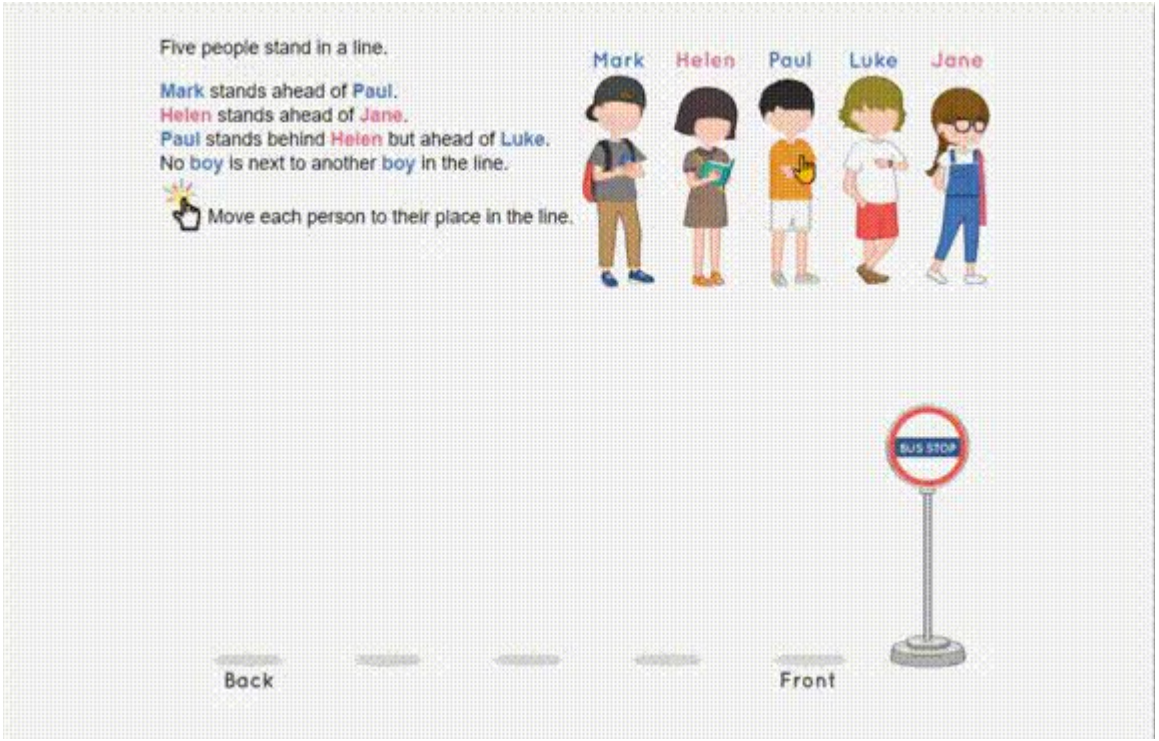


QLens: Visual Analytics of Multi-step Problem-solving Behaviors for Improving Question Design

Meng Xia, Reshika Palaniyappan Velumani, Panpan Xu, Yong Wang, Huamin Qu, Xiaojuan Ma

TVCG 2021

A Multi-step Problem



Motivation

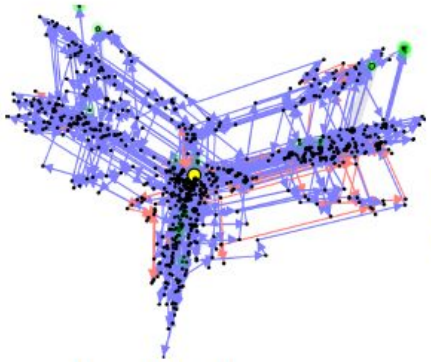
Problem-solving logic

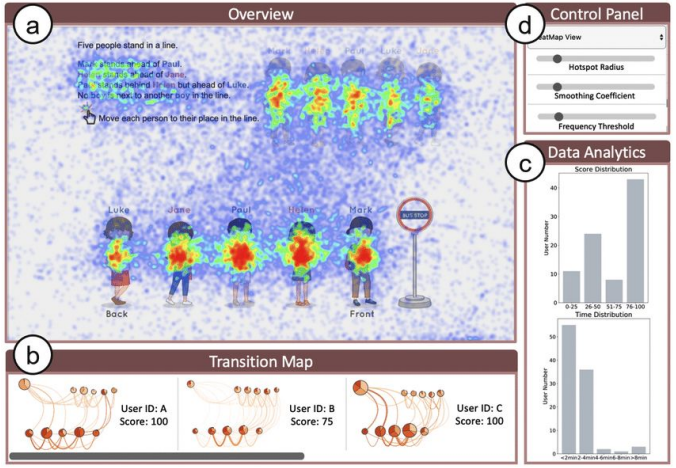
Engagement level

Difficulties

Question Designer

Related work



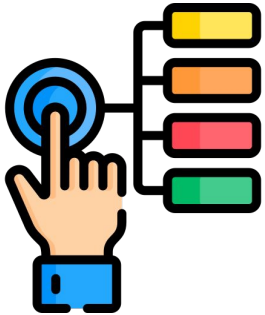


Feature-based projections for effective playtrace analysis (Liu et al., 2011)

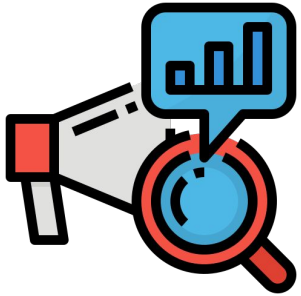
Visual Analytics of Student Learning Behaviors on K-12 Mathematics E-learning Platforms (Xia et al., VIS 2019) ***Best Poster Award***

States cannot reflect students' thinking logic

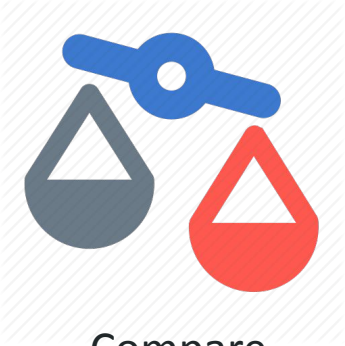
QLens for **question designers**



Inspect



Analyze



Compare

A user-centered design process

- Four domain experts
  - Question designers (E1, E2)
  - System developer (E3)
  - Project manager (E4)
- Requirements gathering iteratively >= one year

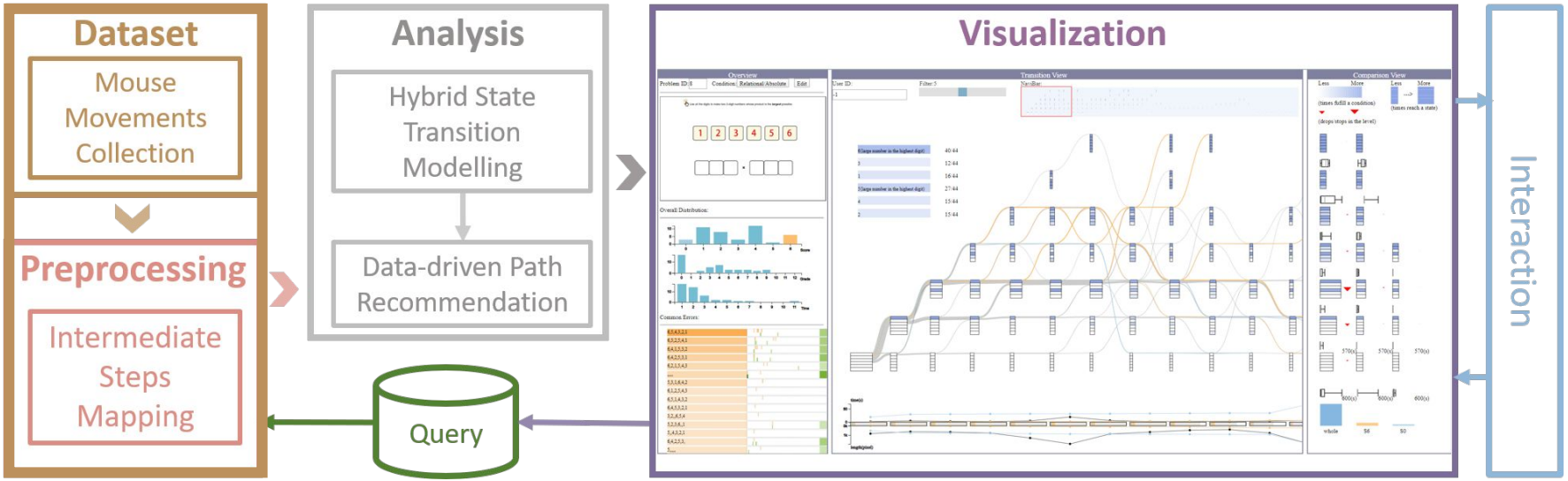
R1: Show students' **overall problem-solving performance**.

R2: **Summarize** and present the multi-step problem-solving behaviors.

R3: **Enable the comparison** of students from different groups.

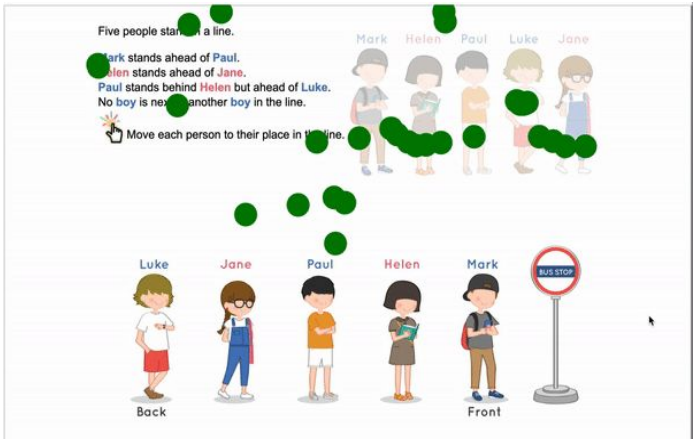
R4: **Evaluate the feasibility** of providing feedback based on existing data.

System overview



1. Data Preprocessing

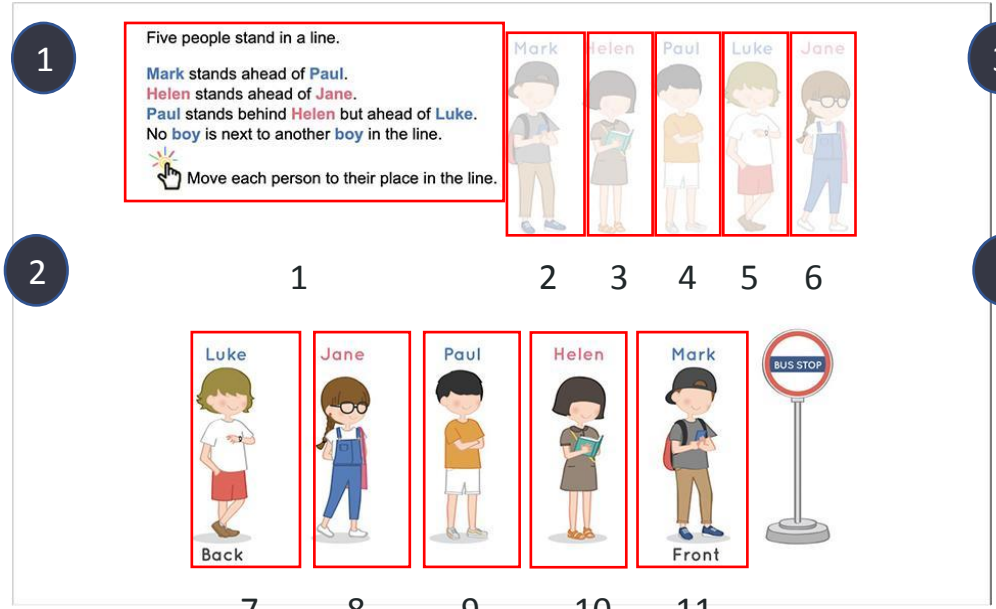
Source URL	http://mad9.learnlex.com/storage/mad/questions/2xbee2fdb4aec4e218/		
Element Path	HTML#.,BODY#.en,DIV#question_content.singlepage,DIV#std_wrapper……		
Question ID	geometry23567	User ID	10001
Time Stamp	20190122T1022	Action Type	click/drag/mousemove
Client Width	1920	Client Height	1080
X	567	Y	432
Touch Screen	True/False	Button	Enter
Platform	Windows/MacOS/iOS	Browser	Chrome/IE/Safari
……			



April 2019 to January 2020,
2,30,644 records from
5,266 students and 1,718 mathematical questions.

1. Data Preprocessing

For each question:



For each student:

3

2 11 4 7 3 8 8 9 …

4

Step1: ,,,,Mark
Step2: Paul,,,,Mark
Step3: Paul,Helen,,,Mark
Step4: Paul,,Helen,,Mark
…

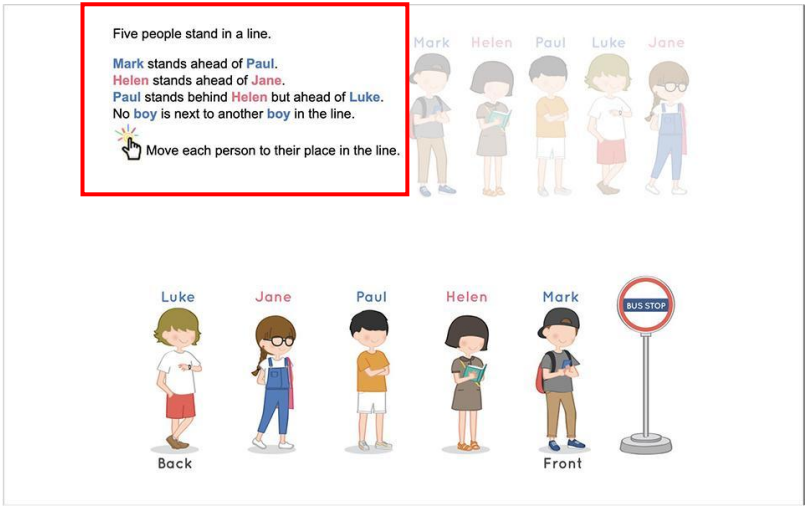
2. Data Analysis - State Transition Model

Step: the smallest user interface interaction that changes the intermediate answers

Stage: the number of conditions the current answer fulfills

Condition: one criteria that students need to fulfill to get the partial score

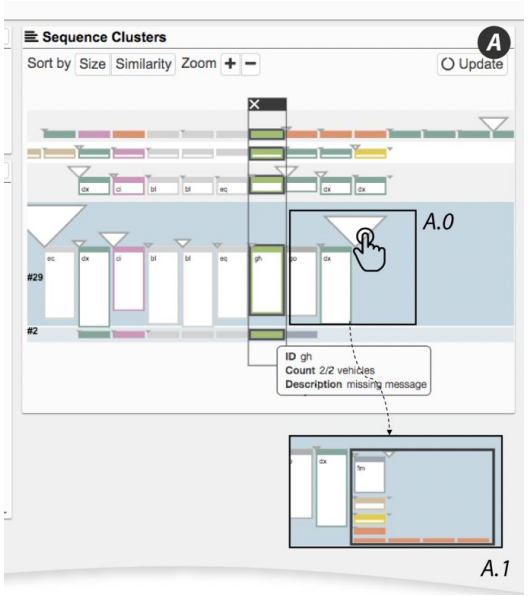
Mark > Paul	149/233
Helen > Jane	140/233
Luke > Paul > Helen	78/233
No boys near each other	0/233



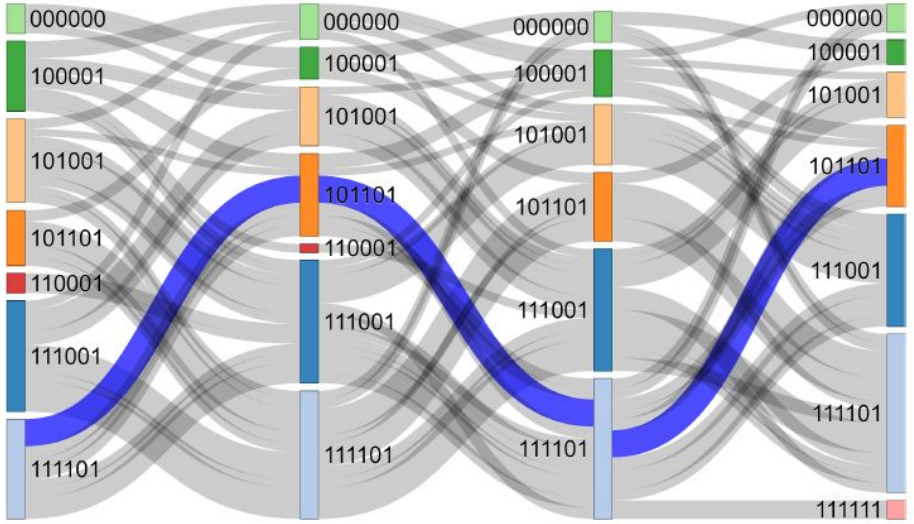
Step1: ,,,,Mark Stage 0
Step2: Paul,,,,Mark Stage 1
Step3: Paul,Helen,,,Mark Stage 2
Step4: Paul,,Helen,,Mark Stage 2
…

3. Visualization - State Transition Visualization

Cannot show event sequence data in which each step has multiple events.

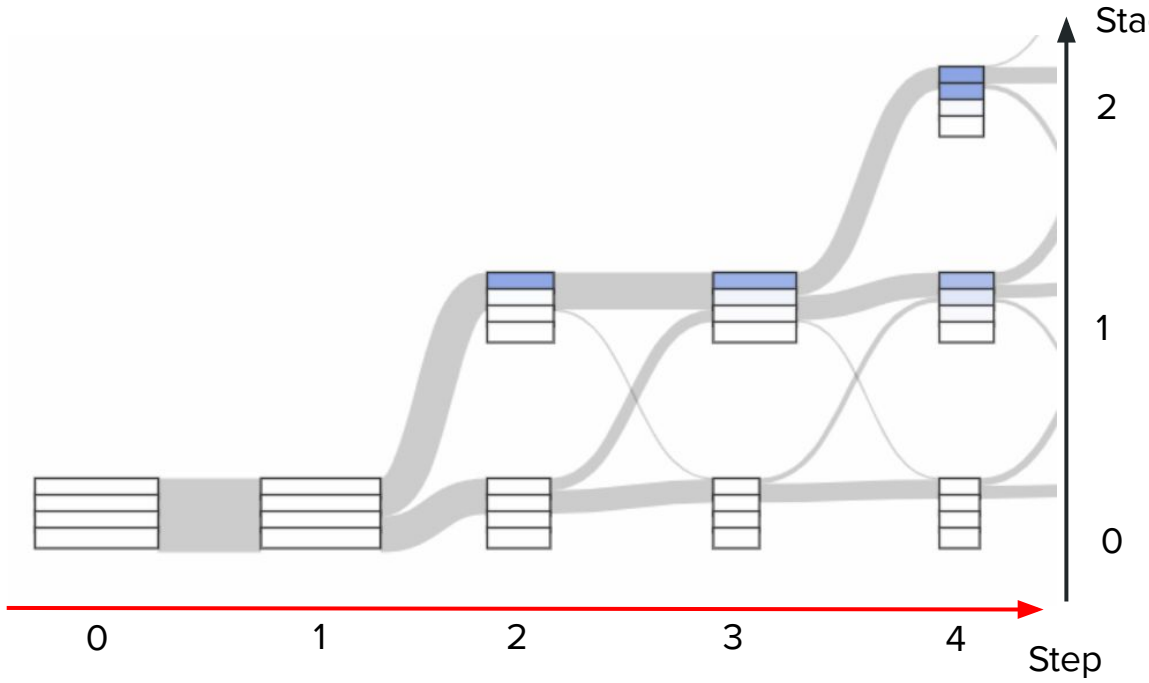


Sequence Synopsis: Optimize Visual Summary of Temporal Event Data (Chen et al., 2017)



Pathviewer: Visualizing pathways through student data (Wang et al., 2017)

3. Visualization - State Transition Visualization

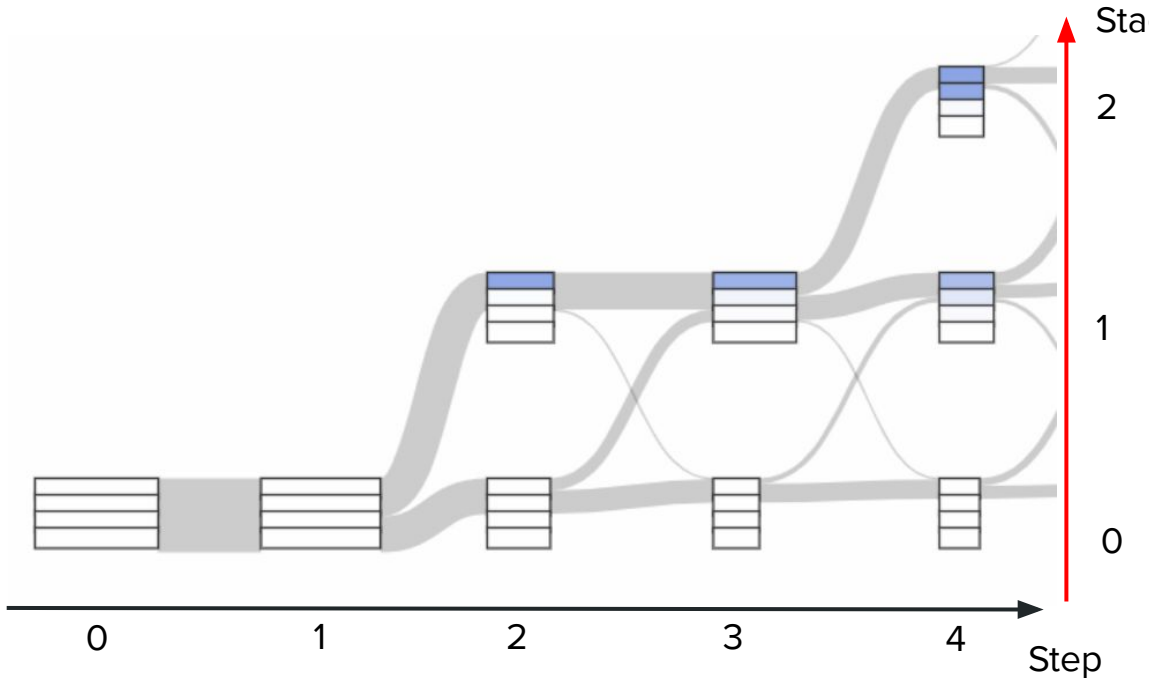


State:

Level1: {Step, Stage} + {Condition array, Time elapse, Trajectory length}

Level2: {Intermediate answer}

3. Visualization - State Transition Visualization

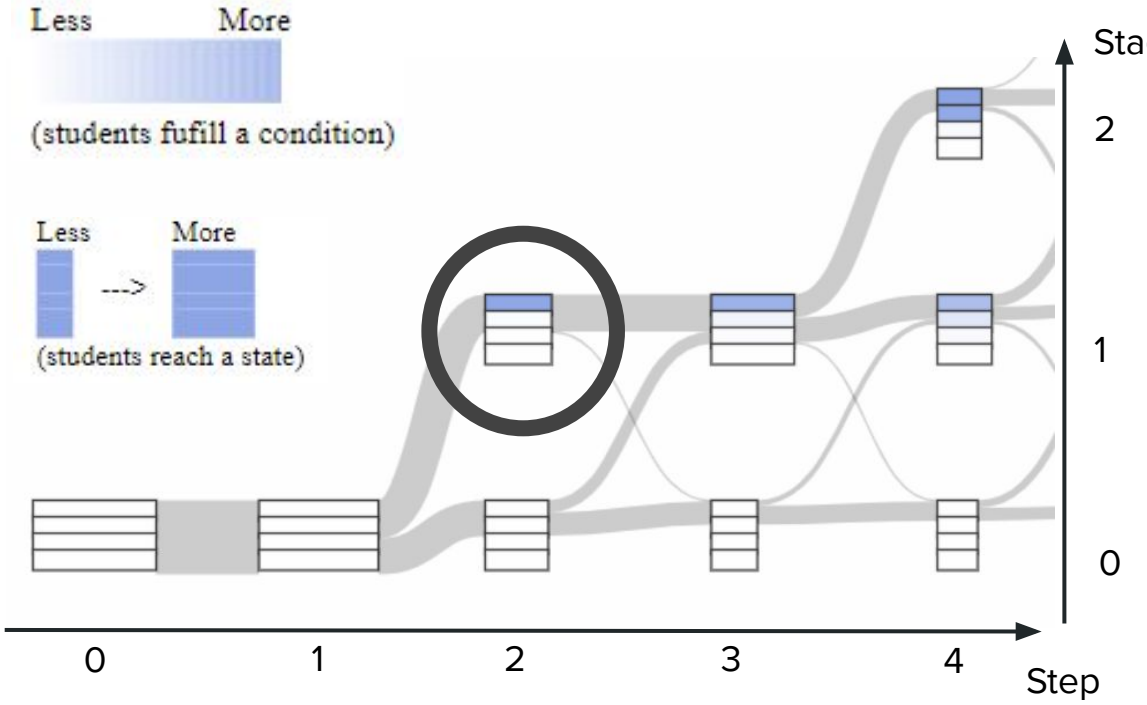


State:

Level1: {Step, Stage} + {Condition array, Time elapse, Trajectory length}

Level2: {Intermediate answer}

3. Visualization - State Transition Visualization

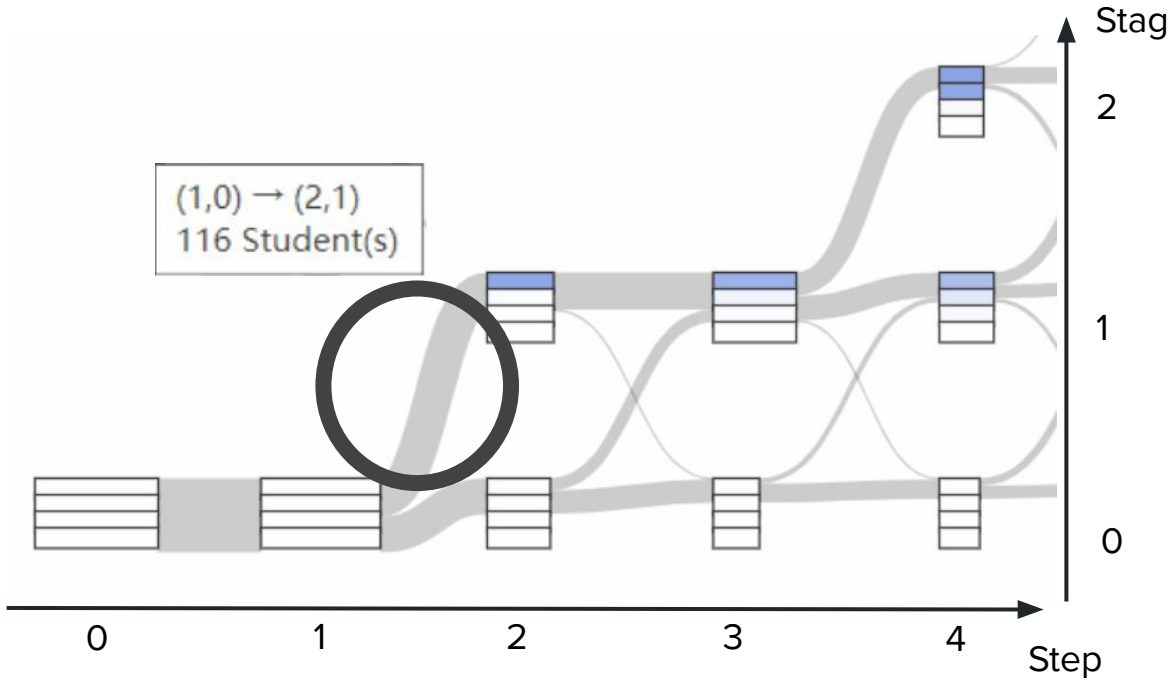


State:

Level1: {Step, Stage} + {Condition array, Time elapse, Trajectory length}

Level2: {Intermediate answer}

3. Visualization - State Transition Visualization

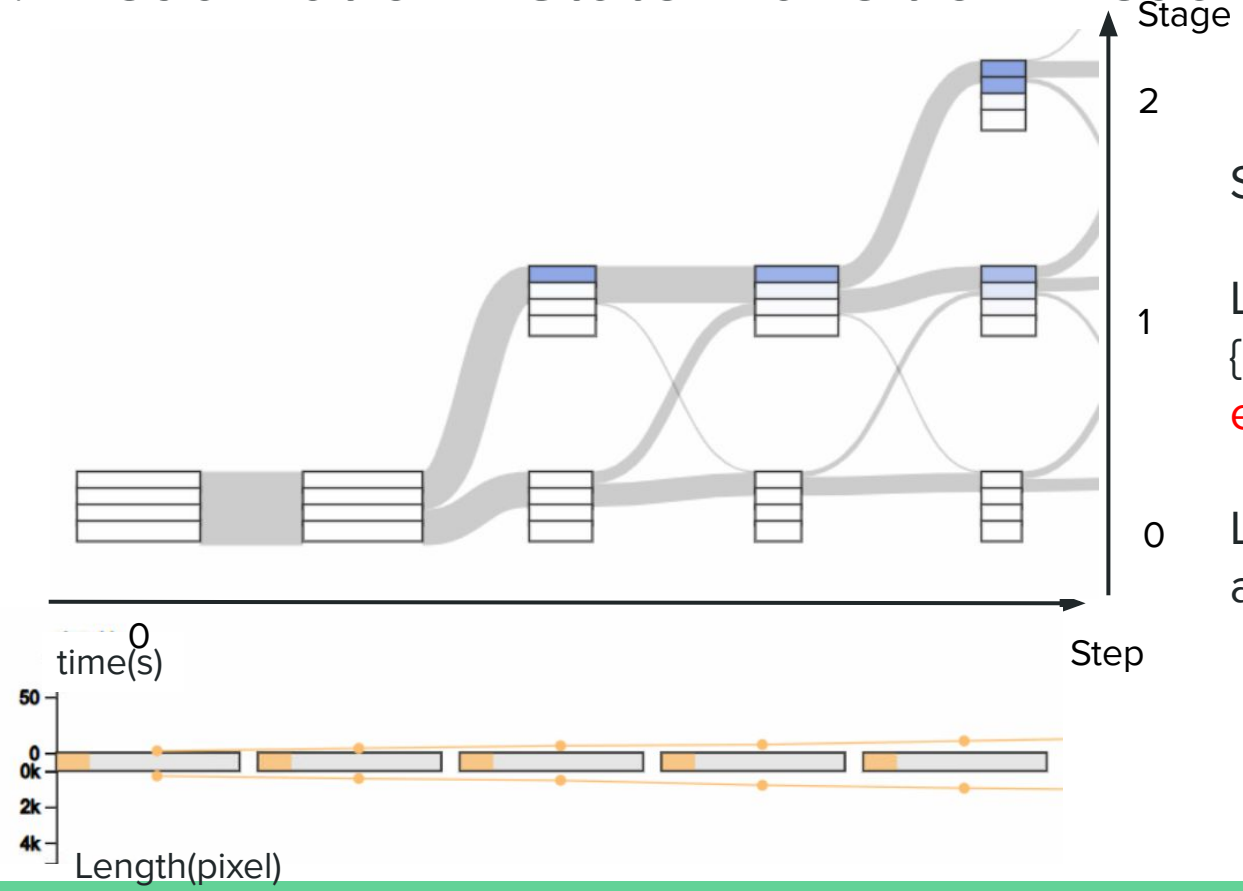


State:

Level1: {Step, Stage} + {Condition array, Time elapse, Trajectory length}

Level2: {Intermediate answer}

3. Visualization - State Transition Visualization

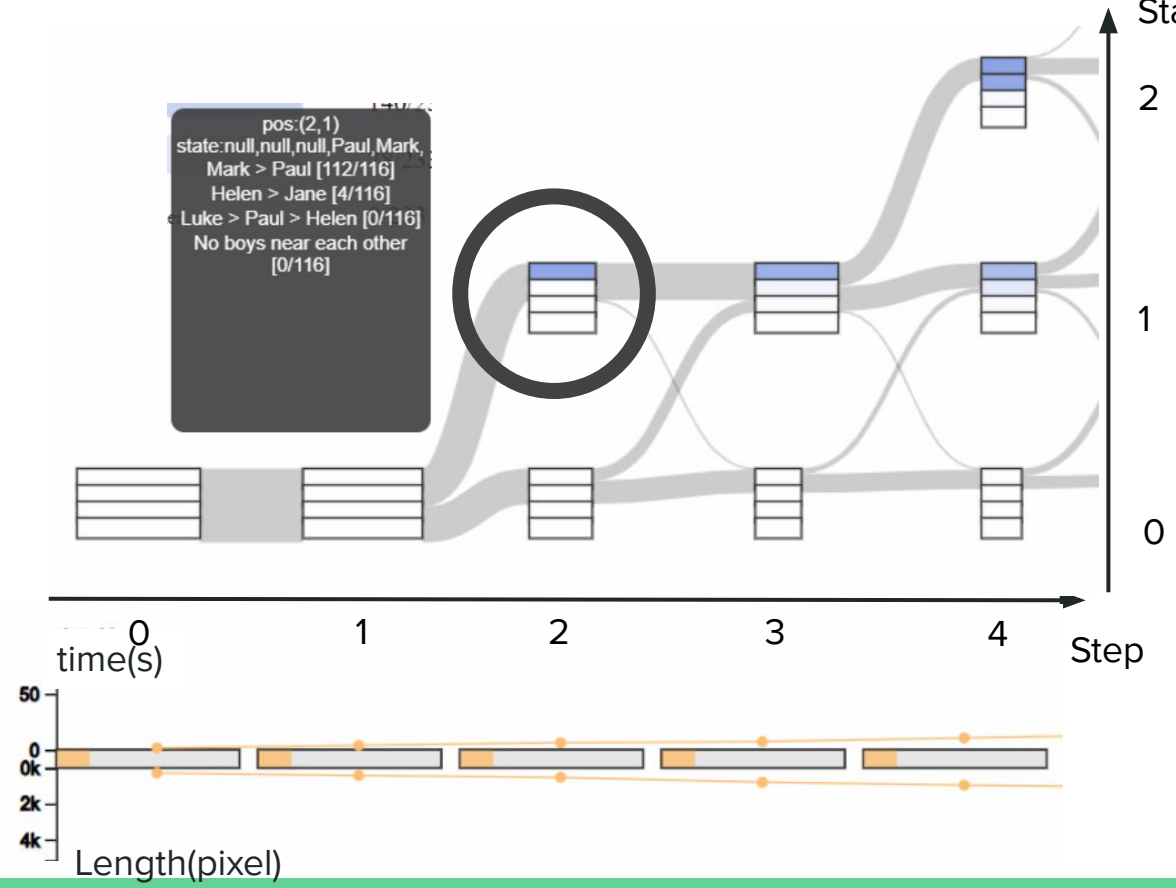


State:

Level1: {Step, Stage} + {Condition array, Time elapse, Trajectory length}

Level2: {Intermediate answer}

3. Visualization - State Transition Visualization

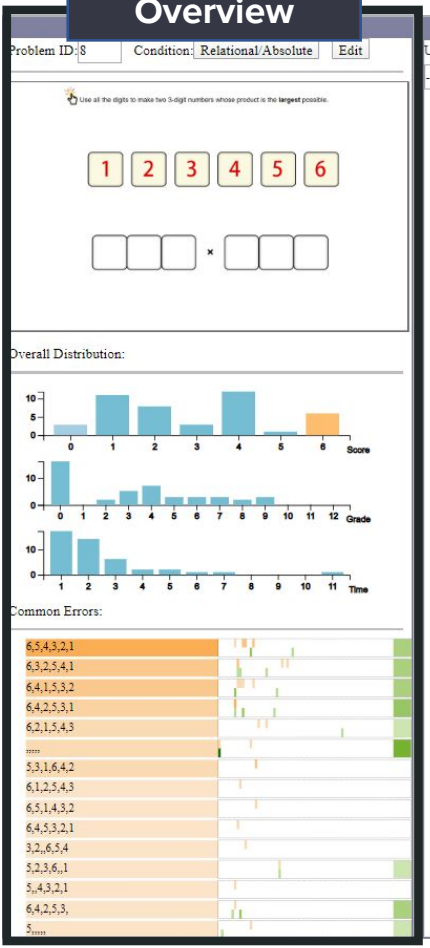


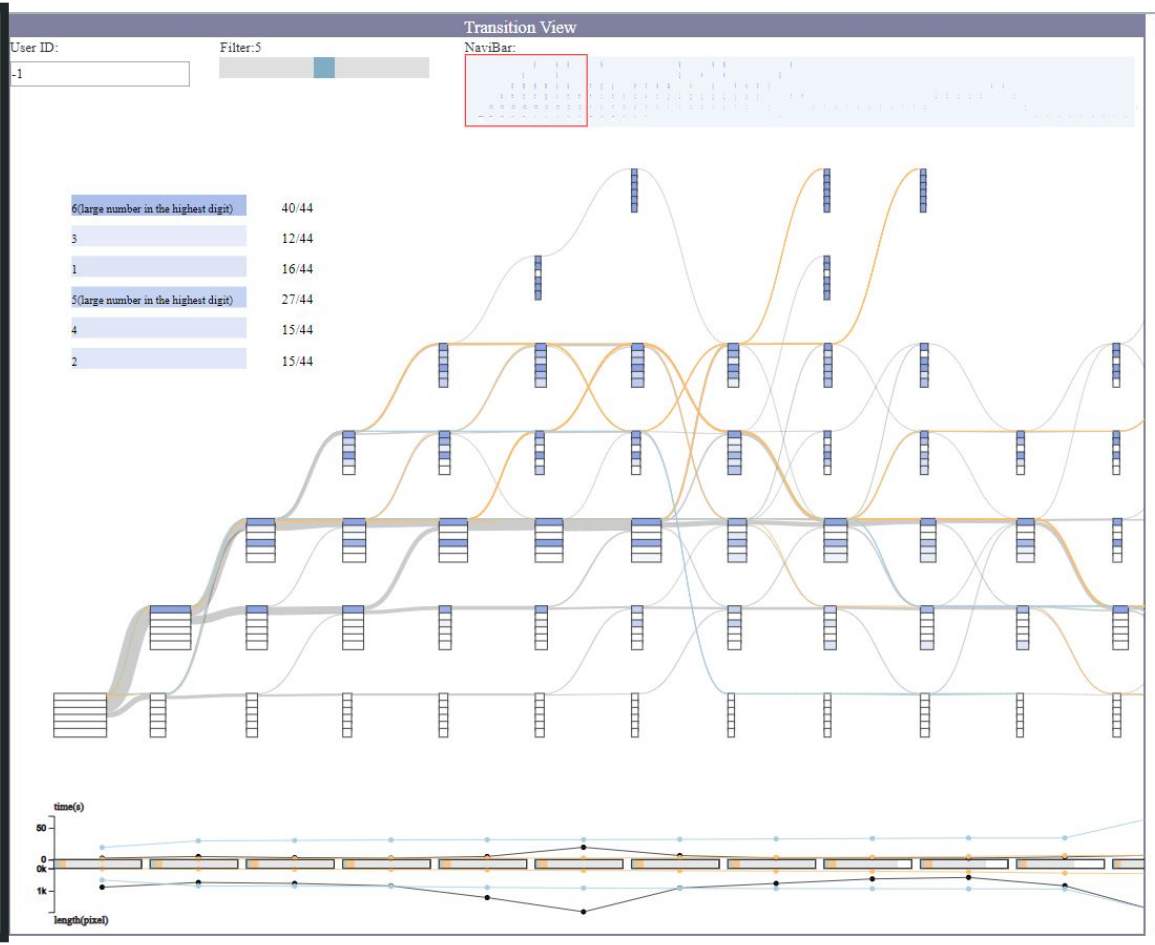
State:

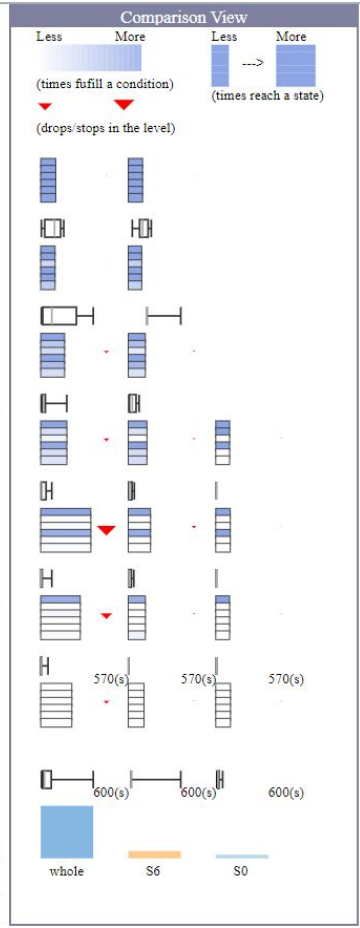
Level1: {Step, Stage} + {Condition array, Time elapse, Trajectory length}

Level2: {Intermediate answer}

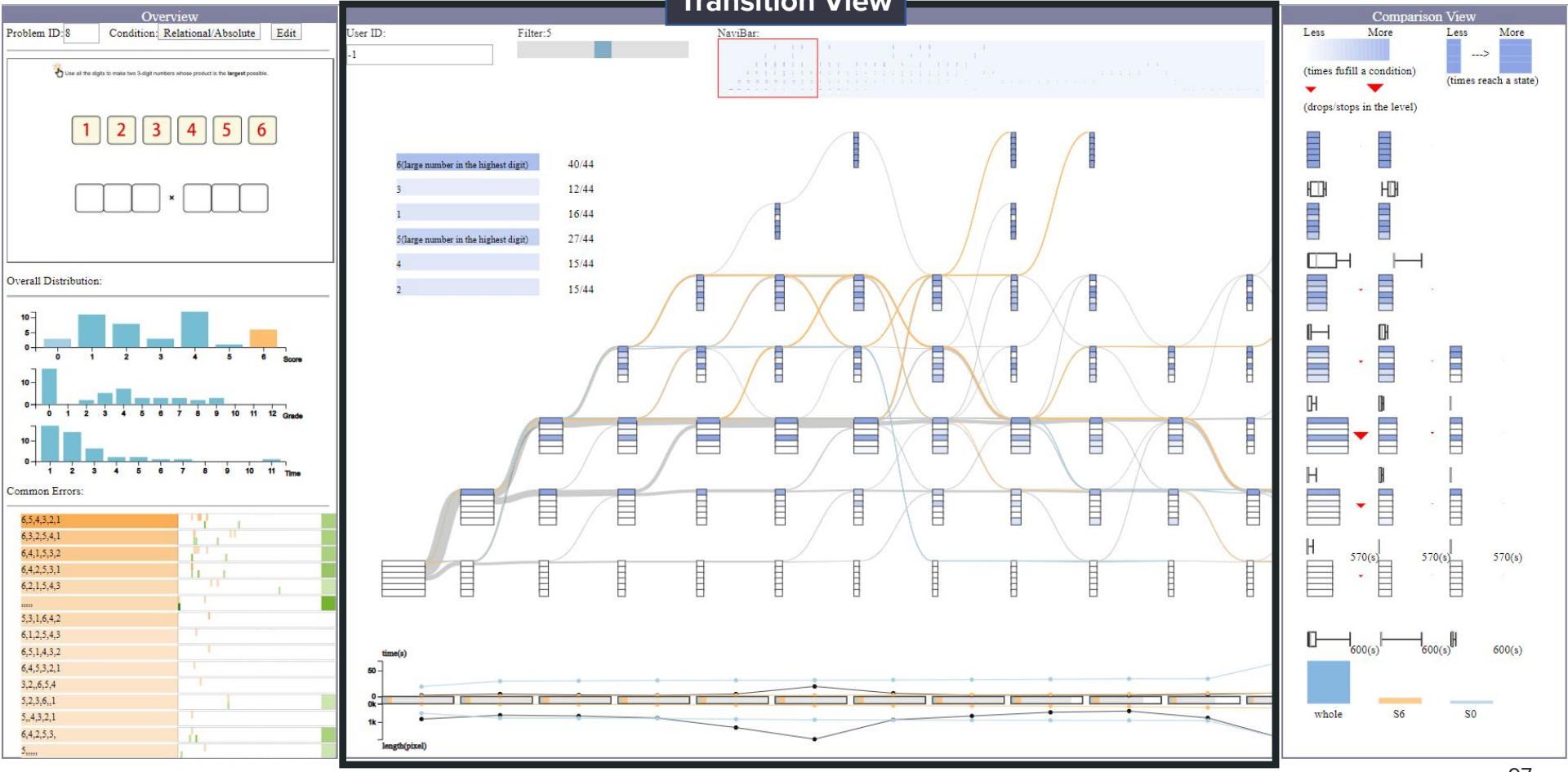
Overview

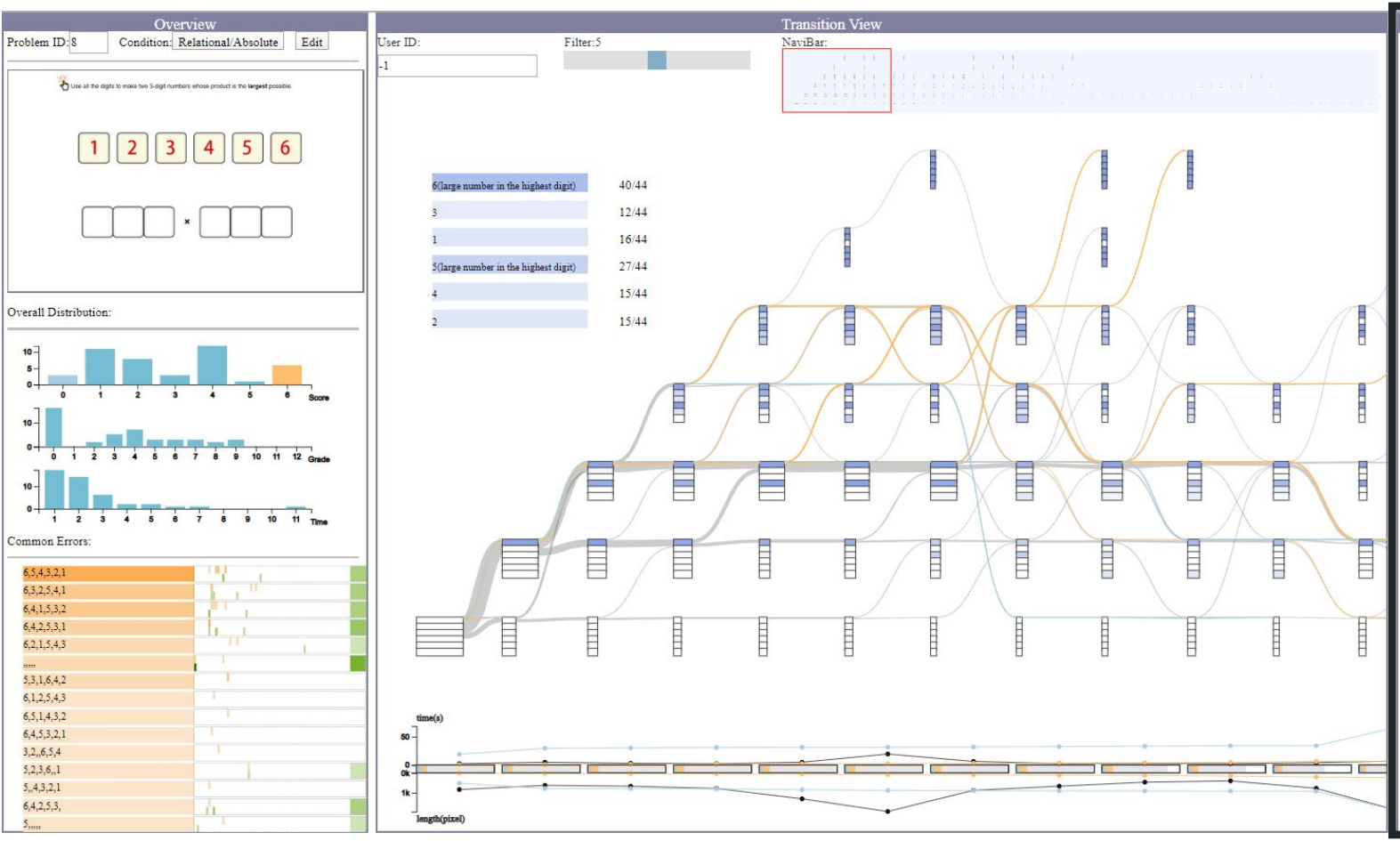

Problem ID: 8
Condition: Relational/Absolute
Edit
Use all the digits to make two 3-digit numbers whose product is the largest possible.
1 2 3 4 5 6
Overall Distribution:
Score
Grade
Time
Common Errors:
6,5,4,3,2,1
6,3,2,5,4,1
6,4,1,5,3,2
6,4,2,5,3,1
6,2,1,5,4,3
5,3,1,6,4,2
6,1,2,5,4,3
6,5,1,4,3,2
6,4,5,3,2,1
3,2,,6,5,4
5,2,3,6,,1
5,,4,3,2,1
6,4,2,5,3,
5,,,,,

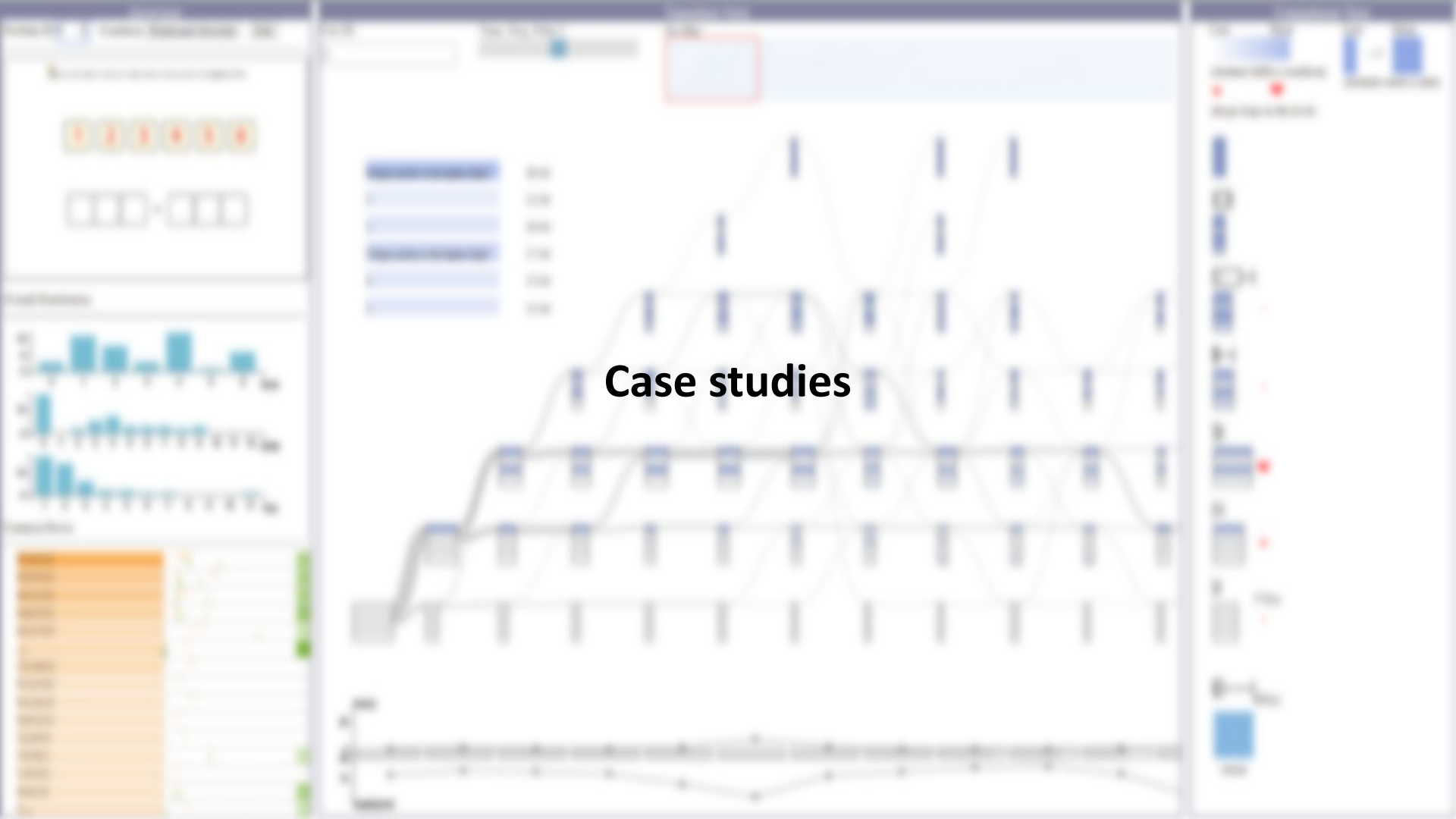

Transition View
User ID:
-1
Filter:5
NaviBar:
6(large number in the highest digit) 40/44
3 12/44
1 16/44
5(large number in the highest digit) 27/44
4 15/44
2 15/44
time(s)
length(pixel)


Comparison View
Less
More
(times fufill a condition)
Less
More
(times reach a state)
(drops/stops in the level)
570(s)
570(s)
570(s)
600(s)
600(s)
600(s)
whole
S6
S0

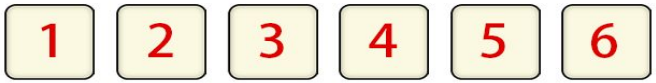
Transition View




Overview
Problem ID: 8
Condition: Relational/Absolute
Edit
Use all the digits to make two 3-digit numbers whose product is the largest possible.
1 2 3 4 5 6
Overall Distribution:
Score
Grade
Time
Common Errors:
6,5,4,3,2,1
6,3,2,5,4,1
6,4,1,5,3,2
6,4,2,5,3,1
6,2,1,5,4,3
5,3,1,6,4,2
6,1,2,5,4,3
6,5,1,4,3,2
6,4,5,3,2,1
3,2,,6,5,4
5,2,3,6,,1
5,,4,3,2,1
6,4,2,5,3,
5,,,,,
Transition View
User ID:
-1
Filter:5
NaviBar:
6(large number in the highest digit) 40/44
3 12/44
1 16/44
5(large number in the highest digit) 27/44
4 15/44
2 15/44
time(s)
length(pixel)
Comparison View
Less
More
(times fufill a condition)
(times reach a state)
(drops/stops in the level)
570(s)
600(s)
whole
S6
S0


Case studies

Use all the digits to make two 3-digit numbers whose product is the largest possible.



RMATION QUESTION SOLUTION SCORE

Correct answer: 6 3 1 × 5 4 2

The larger the multiplicand and the multiplier are, the larger the product is.

To make the largest product, the two largest numbers should be used as the hundreds digit.

6 ☐ ☐ × 5 ☐ ☐

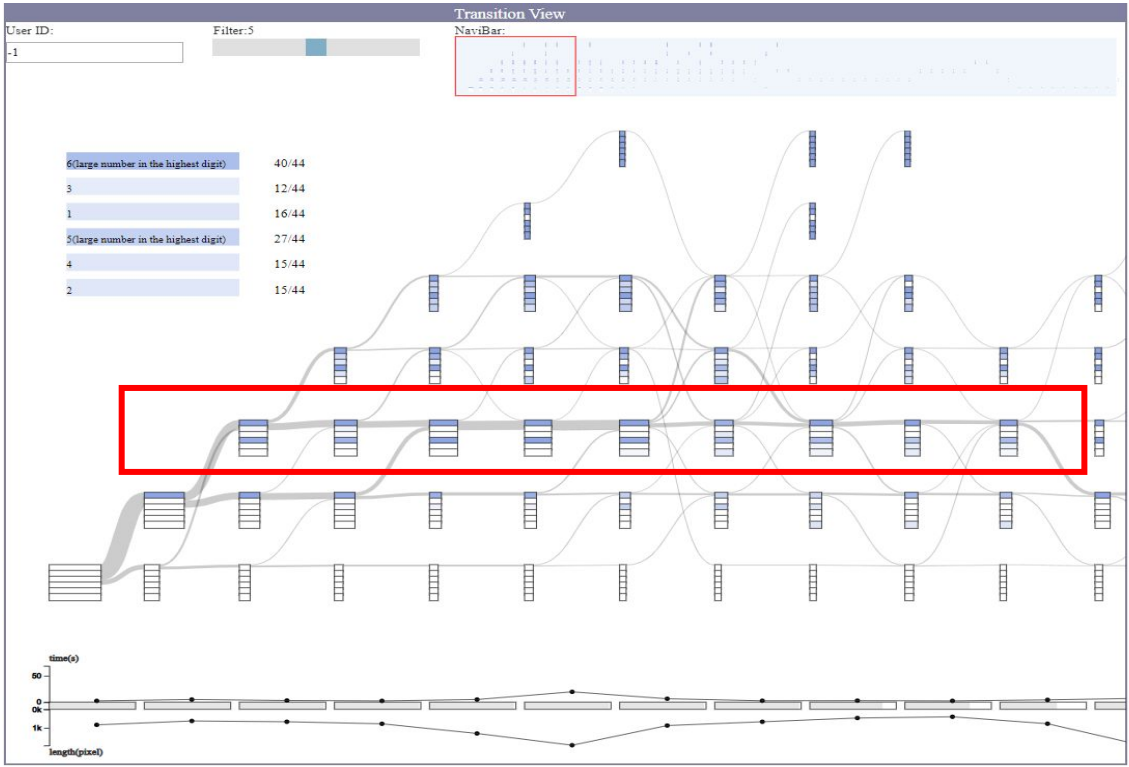
Then the next two largest numbers (3 and 4) should be used for the tens digit.
Consider 2 cases:

Case 1. 6 3 ? × 5 4 ?

63?× 40, must be larger than 24000

Case 2. 6 4 ? × 5 3 ?

64?× 30, must be smaller than 24000



Five people stand in a line.

Mark stands ahead of **Paul**.
Helen stands ahead of **Jane**.
Paul stands behind **Helen** but ahead of **Luke**.
No **boy** is next to another **boy** in the line.

Move each person to their place in the line.

Back

Front

Consider the condition with the most restrictions: "No boy is next to another boy in the line."
There are only 3 boys and 2 girls, so we have

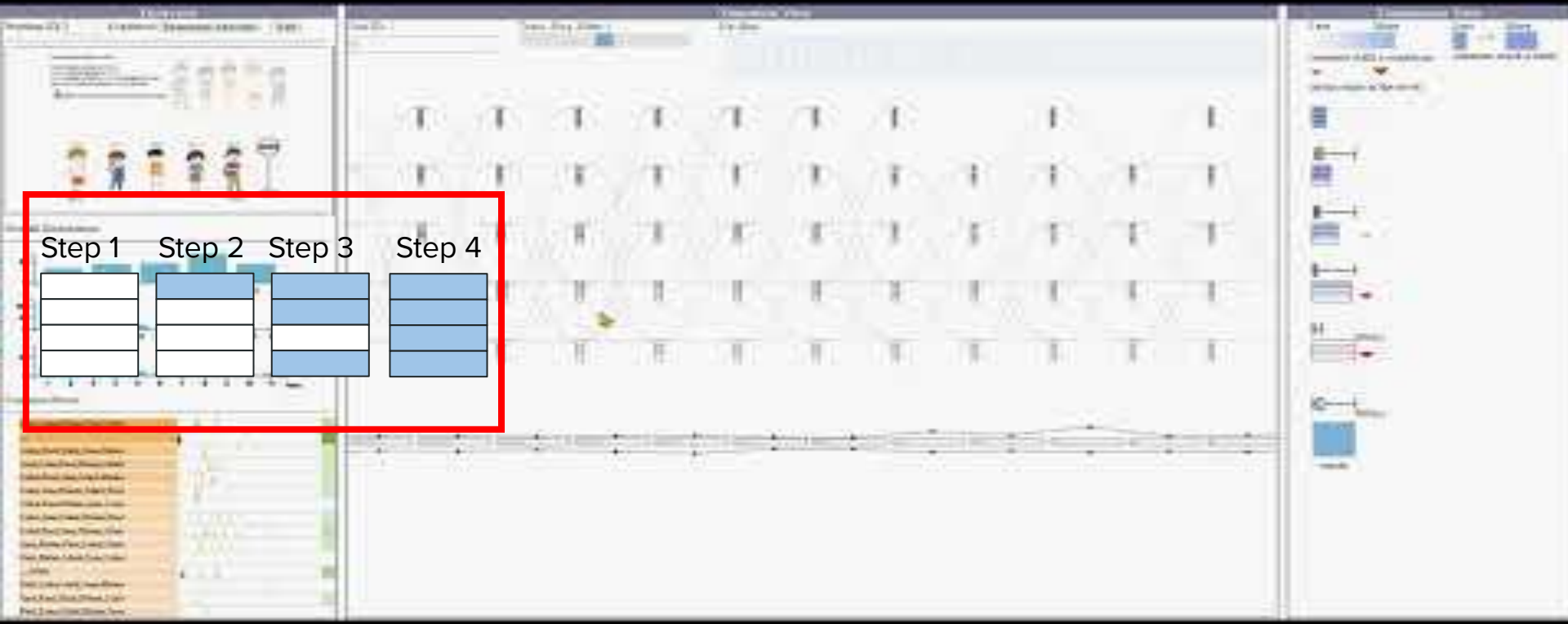
Back Boy Girl Boy Girl Boy Front

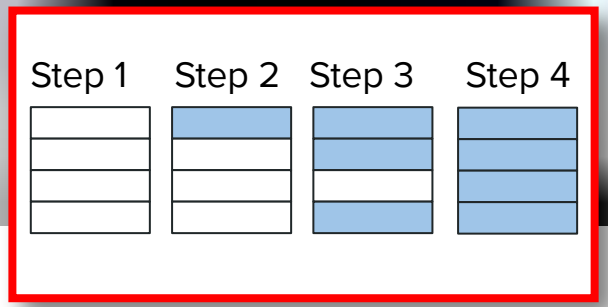
"Mark stands ahead of Paul", "Paul stands ahead of Luke". Therefore,

Back Luke Girl Paul Girl Mark Front

"Helen stands ahead of Jane". Therefore,

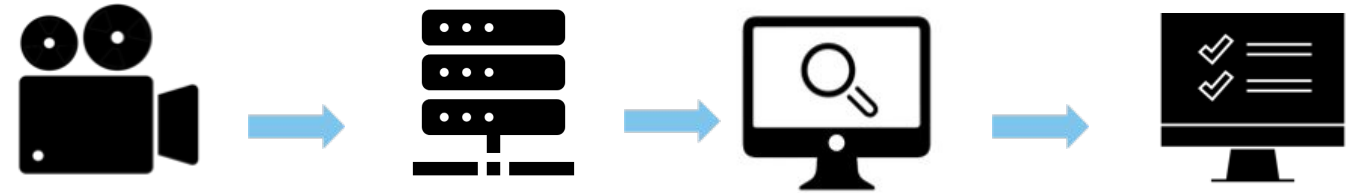
Back Luke Jane Paul Helen Mark Front


Step 1
Step 2
Step 3
Step 4


Step 1
Step 2
Step 3
Step 4

Evaluation

- **Cases studies** with four domain experts during the development
- **Semi-structured interviews** with another three domain experts (two questions designers form a different education company, one senior manager); each interview lasts about 1.5 hours



Evaluation

System usefulness

Overall, all experts confirmed the **usefulness** and the **intuitiveness** of the system.

"The insights from Transition View will be very useful for the question designer (for example to decide which question is more suitable for which grade students) and the system developer."

--- E6

"As more and more learning activities conducted are online, it was also very useful to compare students from different schools (e.g., international and local ones) or regions."

--- E5

"The on-the-fly guidance is what we expected but needs more considerations."

--- E5

Visual design & interactions

"It is so clear to view the problem-solving process using the visualization like this (Transition View)."

--- E7

Positive

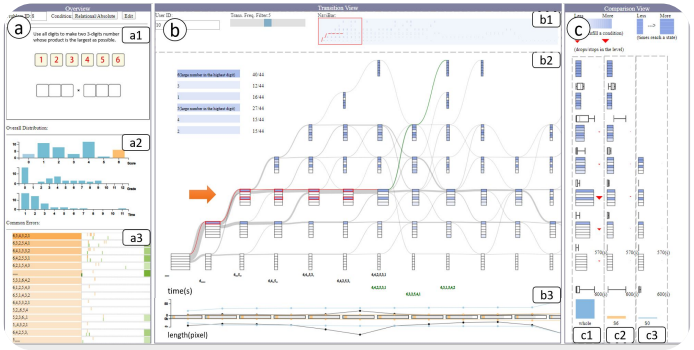
Neutral

Negative

Our collaborator, TrumpTech, uses QLens to improve questions design. The company now serves for **100, 000 students** from more than **500 schools** in Hong Kong.

Conclusion

- An interactive visual analytics system on multi-step question design by analyzing click stream data
- A novel glyph-embedded Sankey diagram for analyzing event sequence trend and comparison, where each step has multiple events
- Three case studies and interviews with domain experts to show the usefulness and usability



My works

Learners (learning loop)

Educators (design loop)

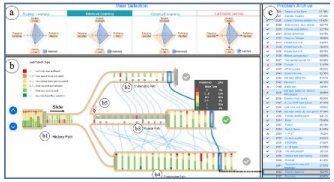
Customizing
learning goals and personalize activities

Reflecting
self-regulations on learning behaviors

Analyzing
learners behaviors and improve learning design

Data

- Large heterogeneous data
- Limited expertise and time in data analysis
- Not enough guidance and explanations
- No guarantee of data quality

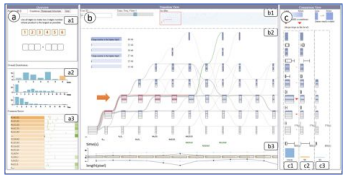


Peerlens (CHI 2019)

- Large heterogeneous data
- Limited expertise and time in data analysis
- Lack of motivation, consistent mental model, and actionable plans

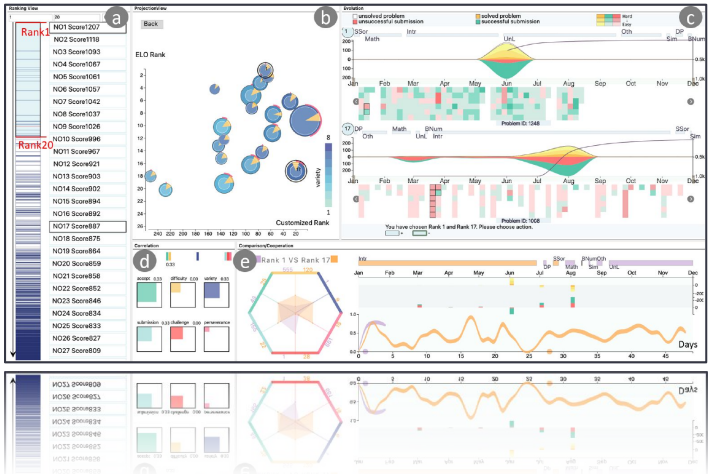
RLens (L@S 2022)

- Large heterogeneous data
- Limited expertise and time in data analysis
- No predefined model



QLens
(TVCG 2021)

SeqDynamics
(EuroVIS 2020)



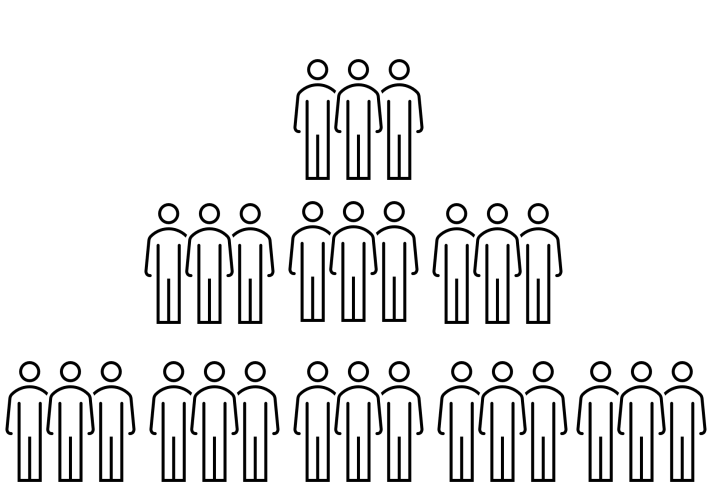
SeqDynamics: Visual Analytics for Evaluating Online Problem-solving Dynamics

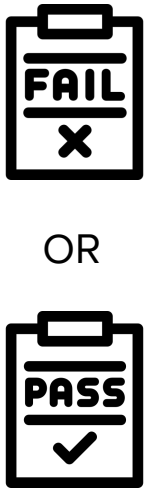
Meng Xia, Min Xu, Chuan-en Lin, Ta Ying Cheng, Huamin Qu, Xiaojuan Ma

EuroVis 2020

Elite Selection in University

Interview in IT Company



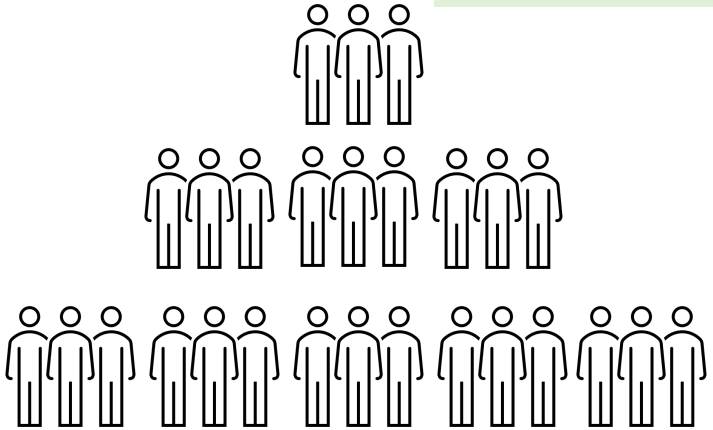


Elite Selection in University

Interview in IT Company

Cognitive skills (think, read, learn, remember, reason, and pay attention)

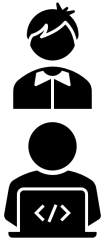
Noncognitive traits (motivation, conscientiousness, perseverance, self-regulation, and collaboration)



FAIL

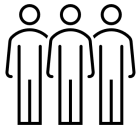
OR

PASS



Elite Selection in University

Interview in IT Company



Cognitive skills (think, read, **learn**, remember, reason, and pay attention)

Noncognitive traits (motivation, conscientiousness, perseverance, **self-regulation**, and collaboration)

OR

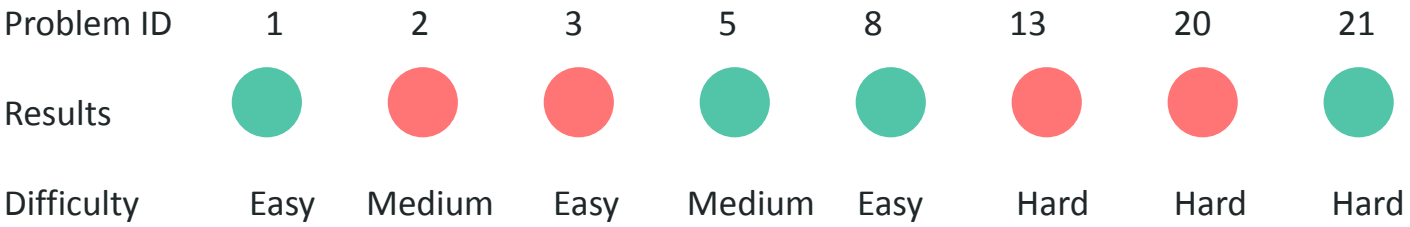
Exams/Technical interviews

Performance and behavior on a long period.

Run ID	Submit Time	Judge Status	Pro.ID	Exe.Time	Exe.Memory	Code Len.	Language	Author
23412857	2017-12-28 00:03:33	Accepted	2046	0MS	1700K	310B	G++	xiameng552180
23412041	2017-12-27 22:24:35	Accepted	2045	0MS	1696K	309B	G++	xiameng552180
23411734	2017-12-27 21:52:45	Wrong Answer	2045	0MS	1700K	388B	G++	xiameng552180
23411669	2017-12-27 21:45:25	Wrong Answer	2045	0MS	1696K	382B	G++	xiameng552180
23411286	2017-12-27 21:10:04	Accepted	2044	15MS	2052K	410B	G++	xiameng552180
23411278	2017-12-27 21:09:05	Wrong Answer	2044	0MS	2048K	404B	G++	xiameng552180
23410918	2017-12-27 20:38:17	Wrong Answer	2044	0MS	2052K	404B	G++	xiameng552180
23408905	2017-12-27 17:39:49	Accepted	2043	0MS	1700K	899B	G++	xiameng552180
23408506	2017-12-27 16:52:47	Accepted	2042	0MS	1696K	254B	G++	xiameng552180
23405316	2017-12-27 01:33:38	Accepted	2041	15MS	1708K	360B	G++	xiameng552180
23405296	2017-12-27 01:21:02	Accepted	2040	78MS	1688K	433B	G++	xiameng552180
23405284	2017-12-27 01:11:27	Accepted	2039	0MS	1740K	280B	G++	xiameng552180
23405283	2017-12-27 01:10:20	Wrong Answer	2039	0MS	1744K	283B	G++	xiameng552180
23405282	2017-12-27 01:09:41	Wrong Answer	2039	15MS	1692K	280B	G++	xiameng552180
23405277	2017-12-27 01:05:09	Accepted	2037	15MS	1708K	729B	G++	xiameng552180

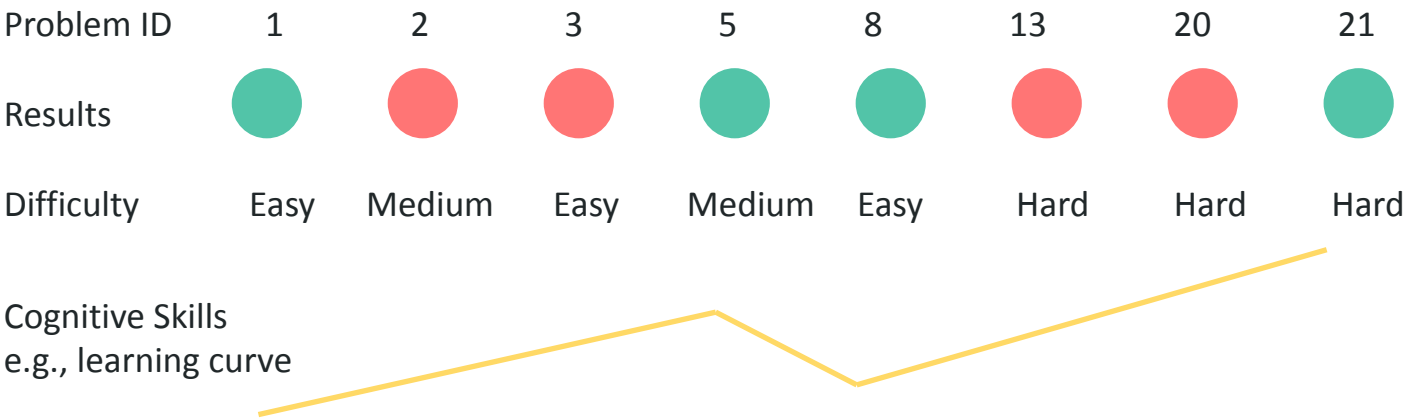
Problem-solving Dynamics

The process and progress of solving a series of problems over time.



Problem-solving Dynamics

The process and progress of solving a series of problems over time.

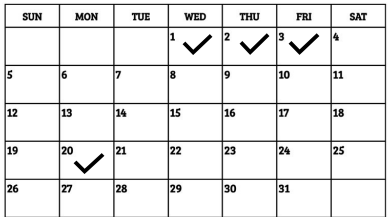


Problem-solving Dynamics

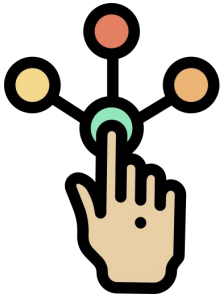
The process and progress of solving a series of problems over time.

Problem ID	1	2	3	5	8	13	20	21
Results								
Difficulty	Easy	Medium	Easy	Medium	Easy	Hard	Hard	Hard
Timestamp	Jan. 1	Jan. 1	Jan. 2	Jan. 3	Jan. 3	Jan. 20	Jan. 20	Jan. 20

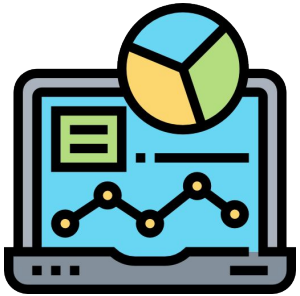
Non-cognitive Trait
e.g., self-regulation



SeqDynamics



Interactive



Multi-dimensional

Time-series

A user-centered design process

Four domain experts

- Recruiters from the competitive programming team (E1, E2)
- Student coaches (E3, E4)

Requirements gathering iteratively for three months

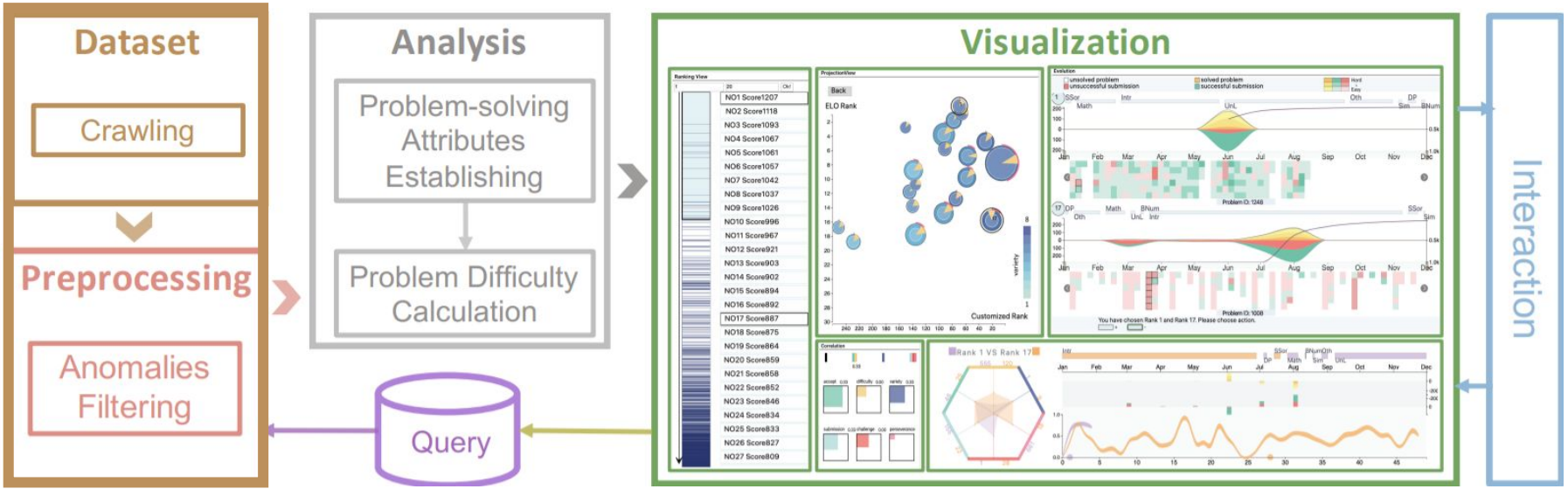
R1: Show a clear overview of overall students' problem-solving performance.

R2: Understand problem-solving dynamics from different perspectives over time. (i.e., cognitive and non-cognitive).

R3: Compare/Combine the problem-solving performance at different scales.

R4: Support an interactive and customized exploration of the evaluation.

System overview



Problem-solving Feature Extraction

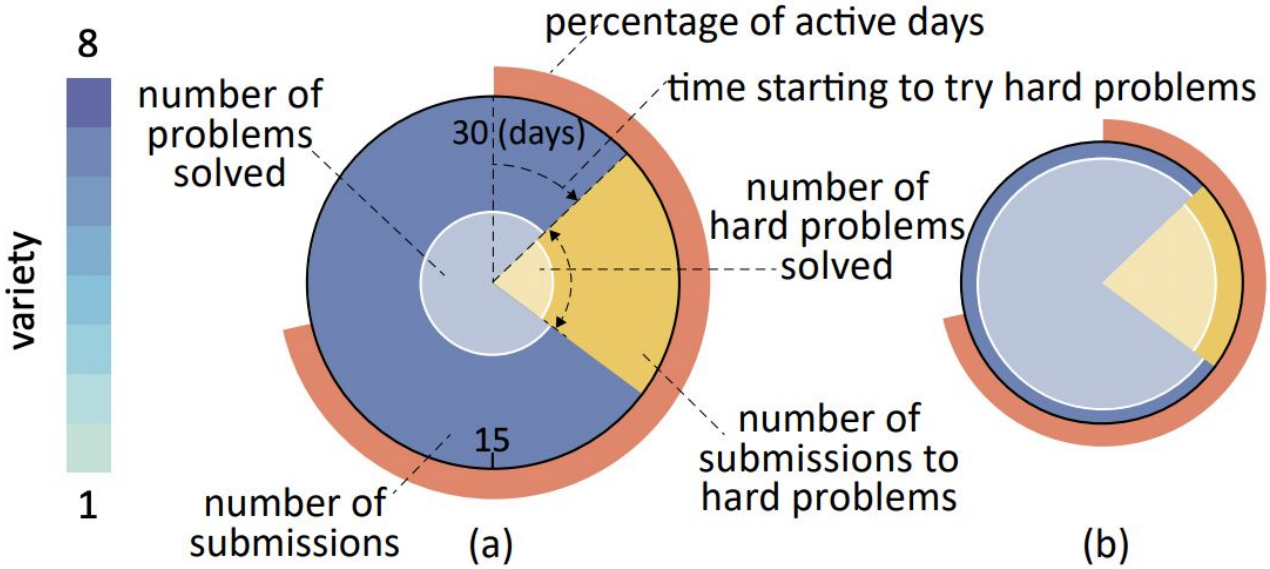
Changes of these features below over time:

Cognitive ability *(Ausubel et al., 1968)*

- L1: number of problems solved
- L2: ratio of hard problems solved
- L3: diversity of problems solved

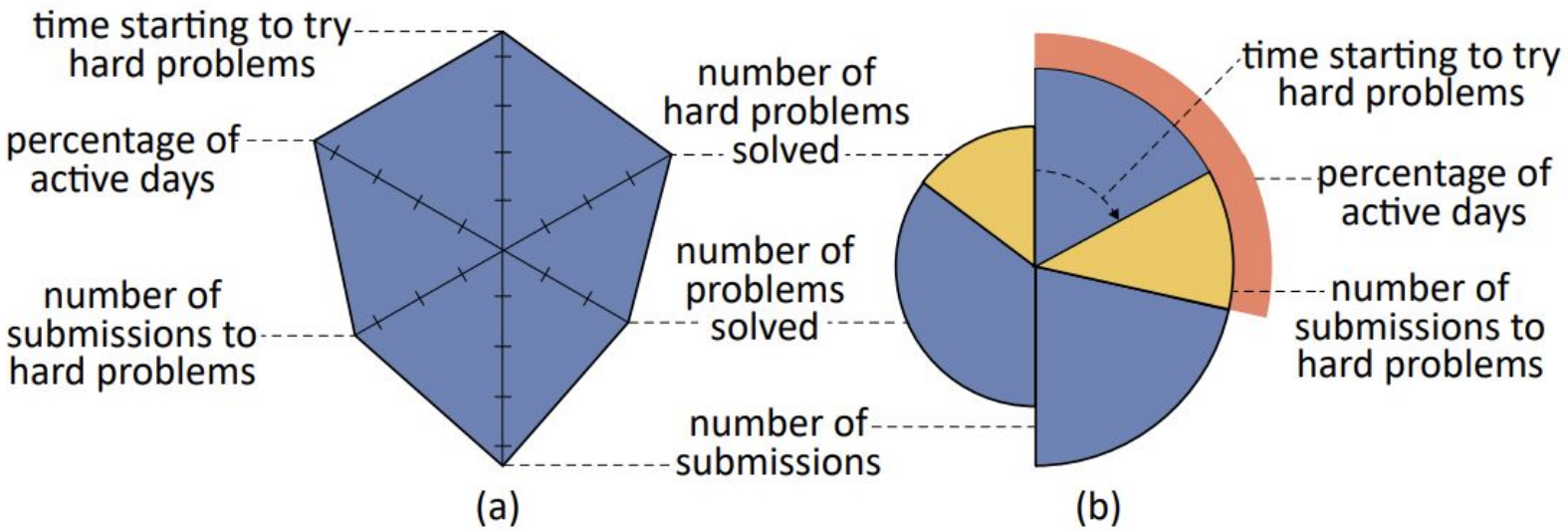
Non-cognitive traits *(Farkas, 2003)*

- L4: number of submissions (diligence level)
- L5: time starting to trying hard problems (willingness to take challenge)
- L6: ratio of active days (perseverance)



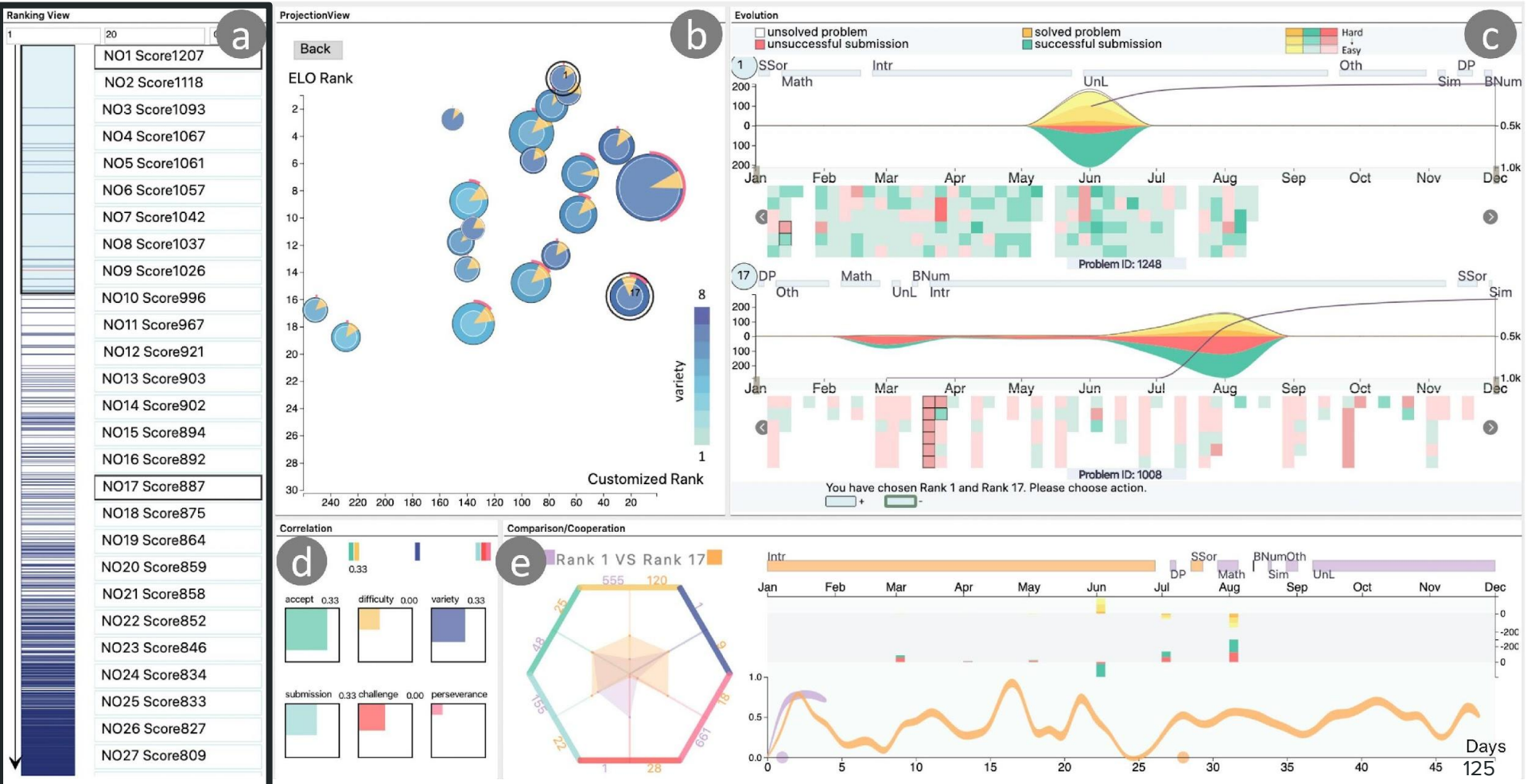
(a) A learner who has many submissions but solves a few problems.

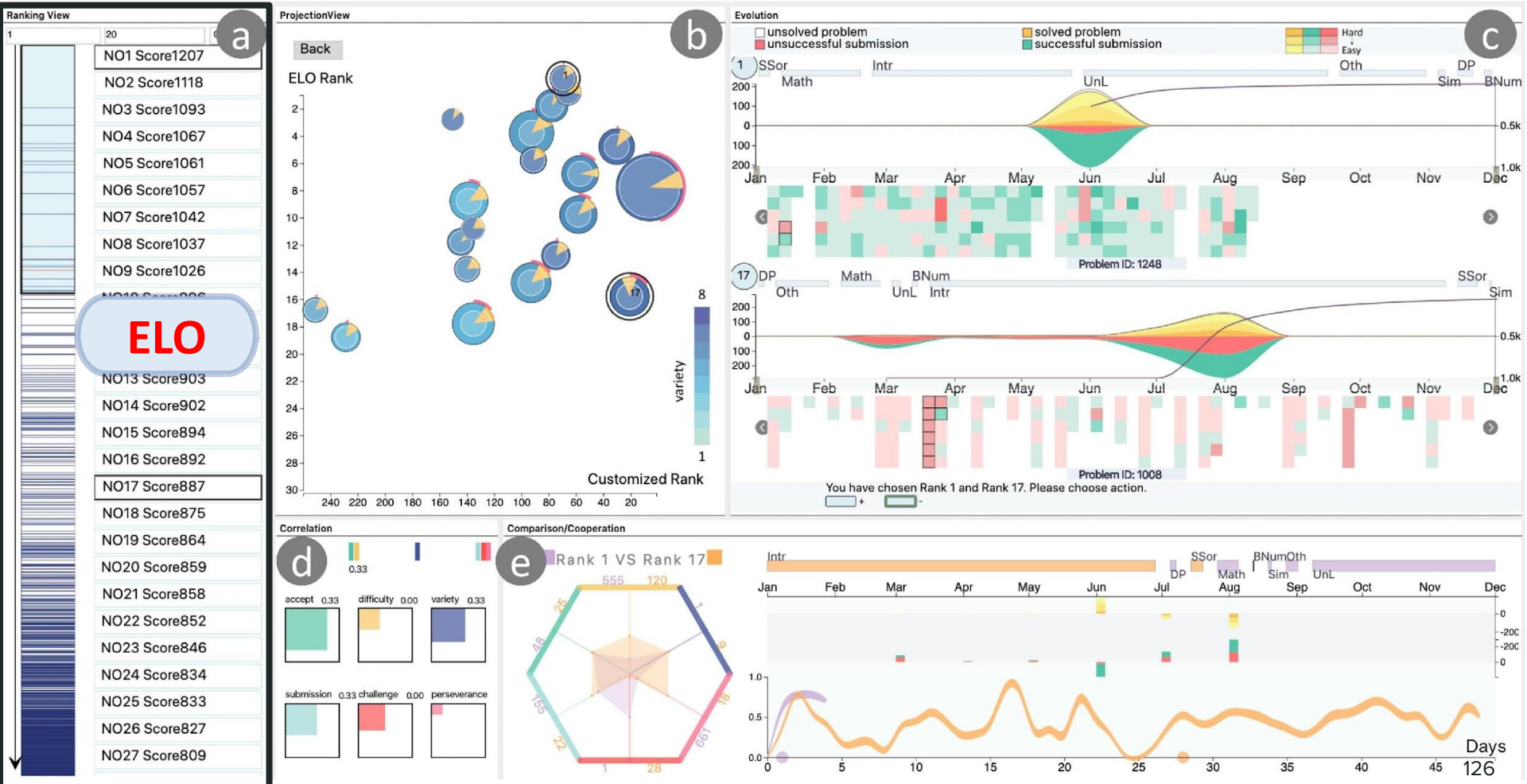
(b) A learner who has relatively fewer submissions but solves more problems.

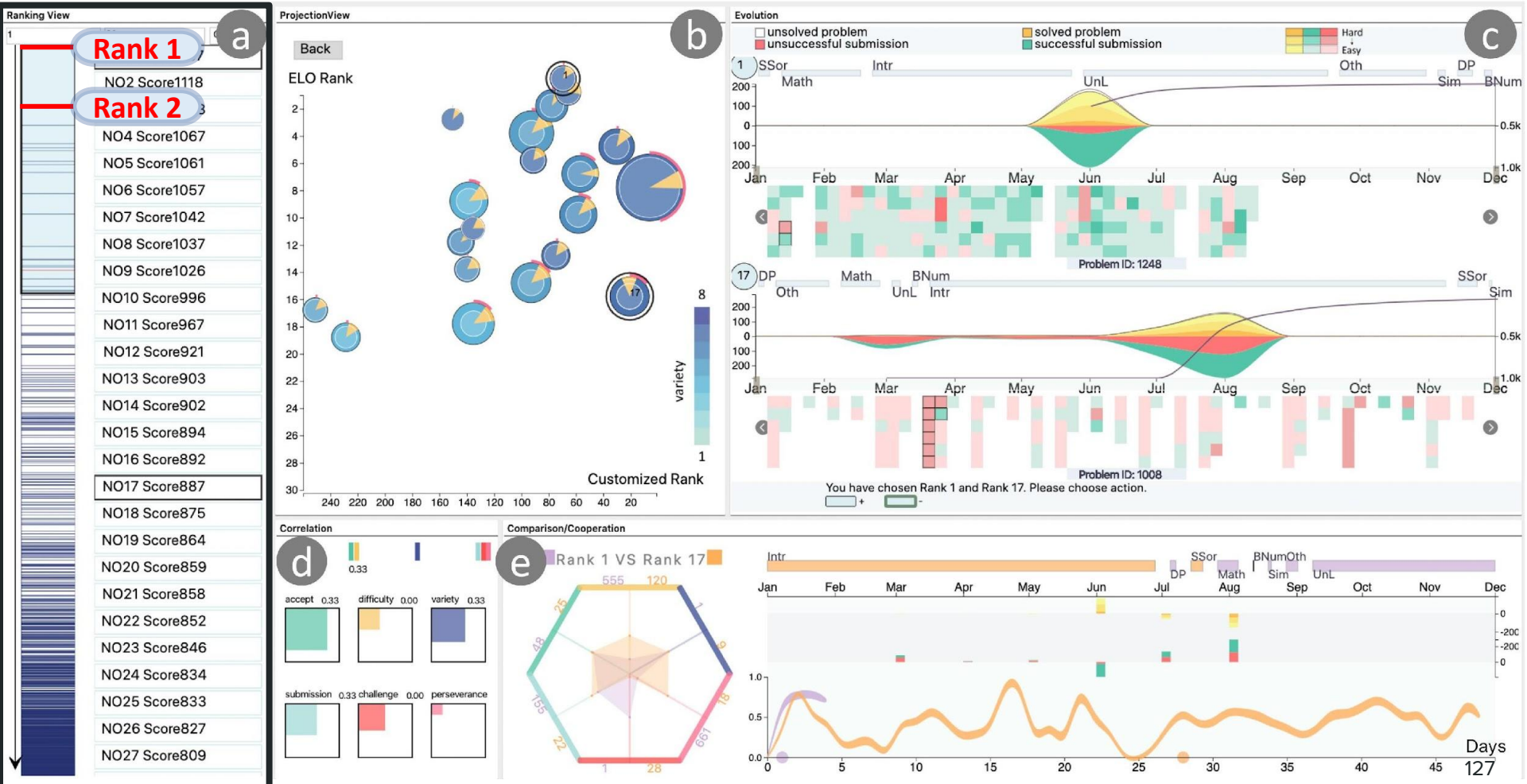


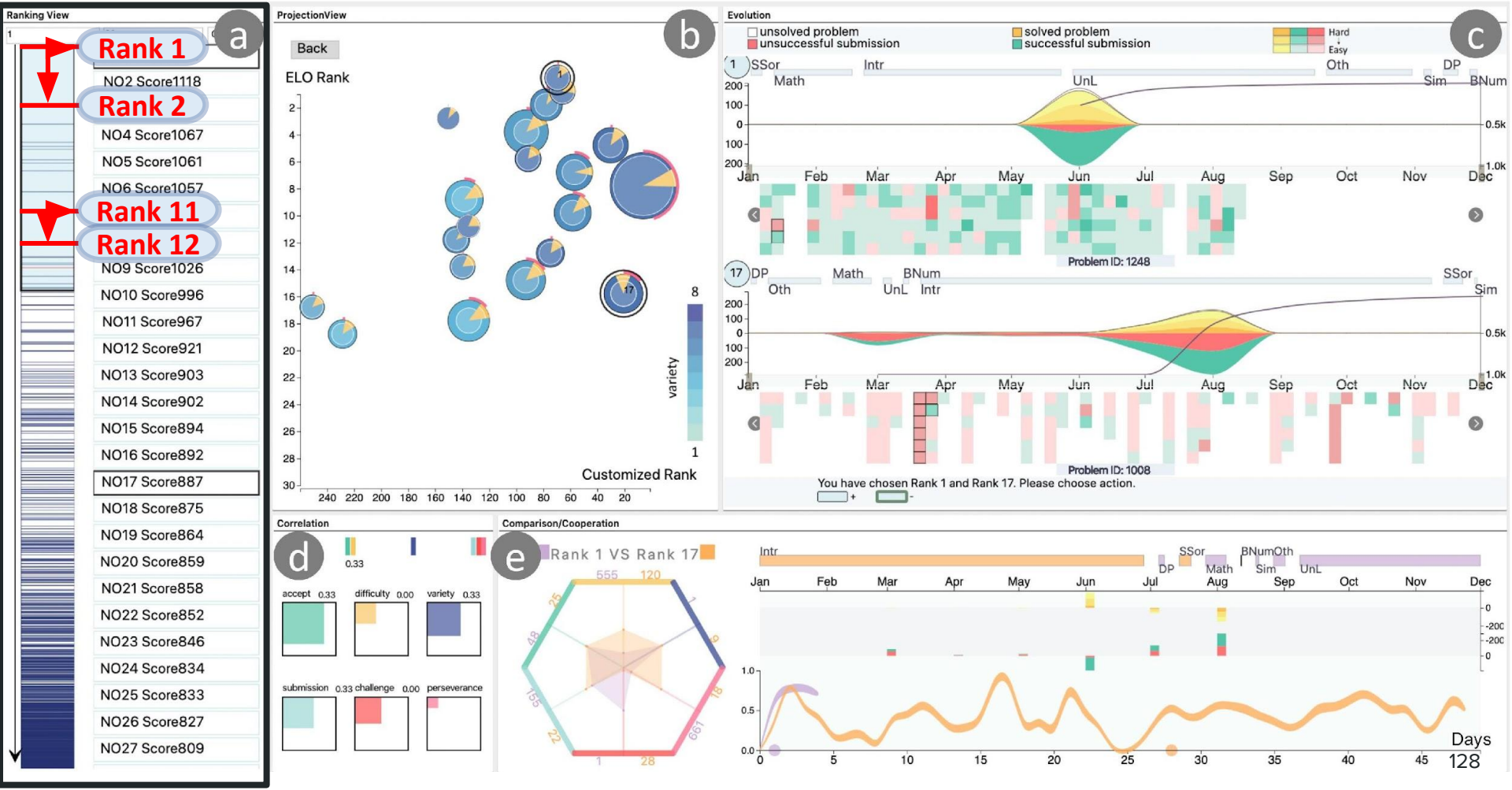
Two design alternatives.

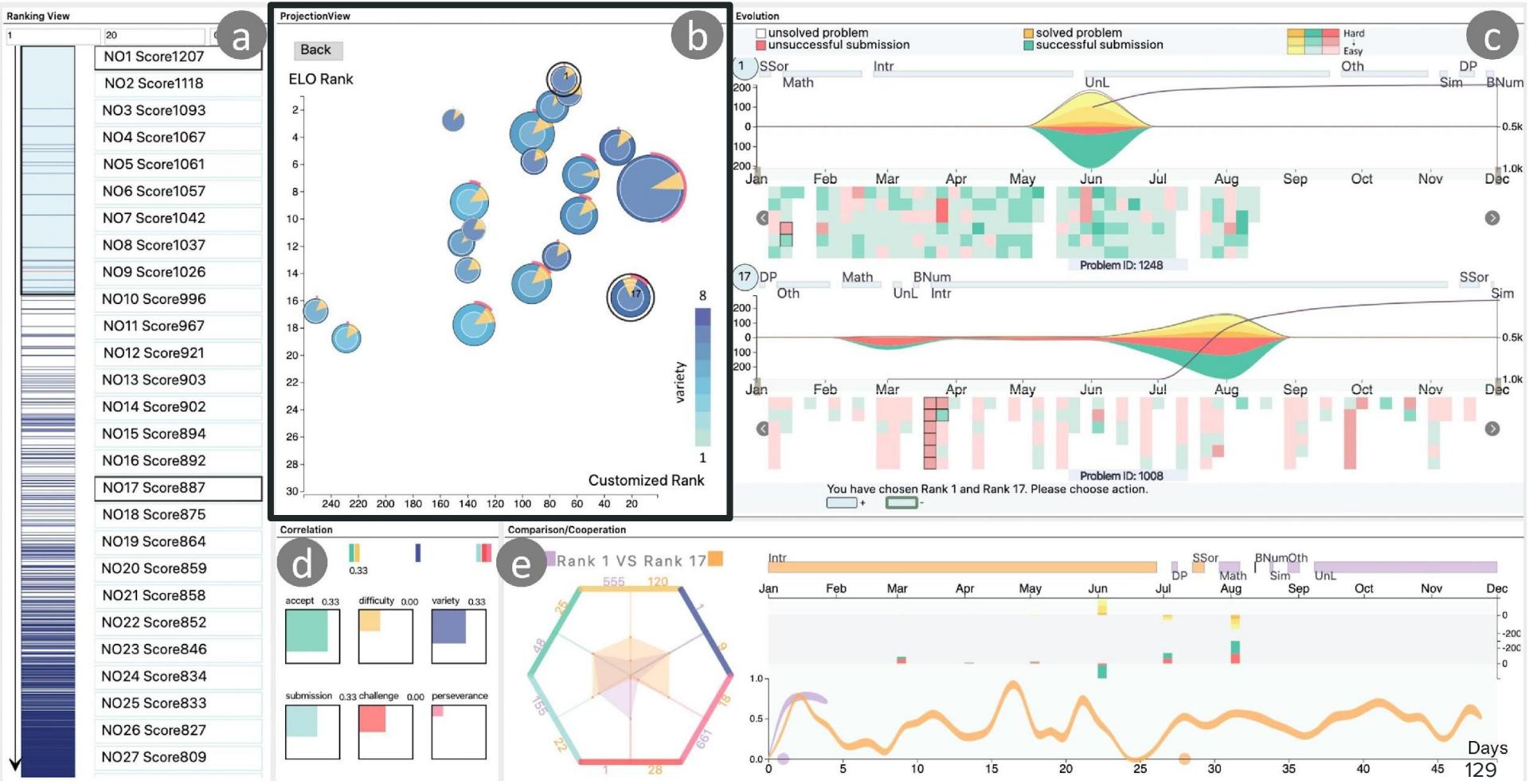
User Interface

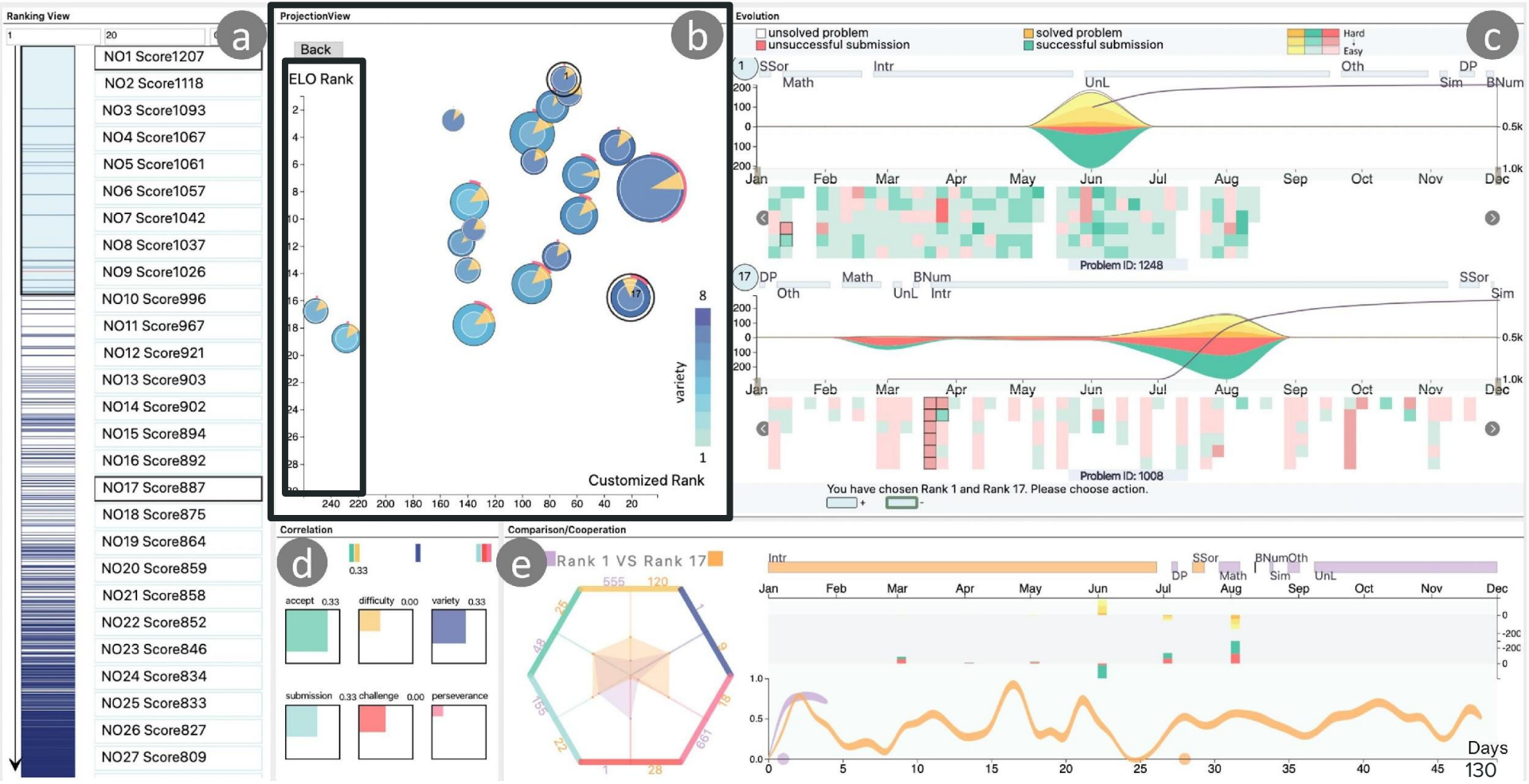

Ranking View
a
NO1 Score1207
NO2 Score1118
NO3 Score1093
NO4 Score1067
NO5 Score1061
NO6 Score1057
NO7 Score1042
NO8 Score1037
NO9 Score1026
NO10 Score996
NO11 Score967
NO12 Score921
NO13 Score903
NO14 Score902
NO15 Score894
NO16 Score892
NO17 Score887
NO18 Score875
NO19 Score864
NO20 Score859
NO21 Score858
NO22 Score852
NO23 Score846
NO24 Score834
NO25 Score833
NO26 Score827
NO27 Score809
ProjectionView
b
Back
ELO Rank
Customized Rank
variety
Evolution
c
unsolved problem
solved problem
unsuccessful submission
successful submission
Hard
Easy
SSor
Intr
Math
UnL
Oth
DP
Sim
BNum
Problem ID: 1248
Problem ID: 1008
You have chosen Rank 1 and Rank 17. Please choose action.
Correlation
d
accept
difficulty
variety
submission
challenge
perseverance
Comparison/Cooperation
e
Rank 1 VS Rank 17
Days

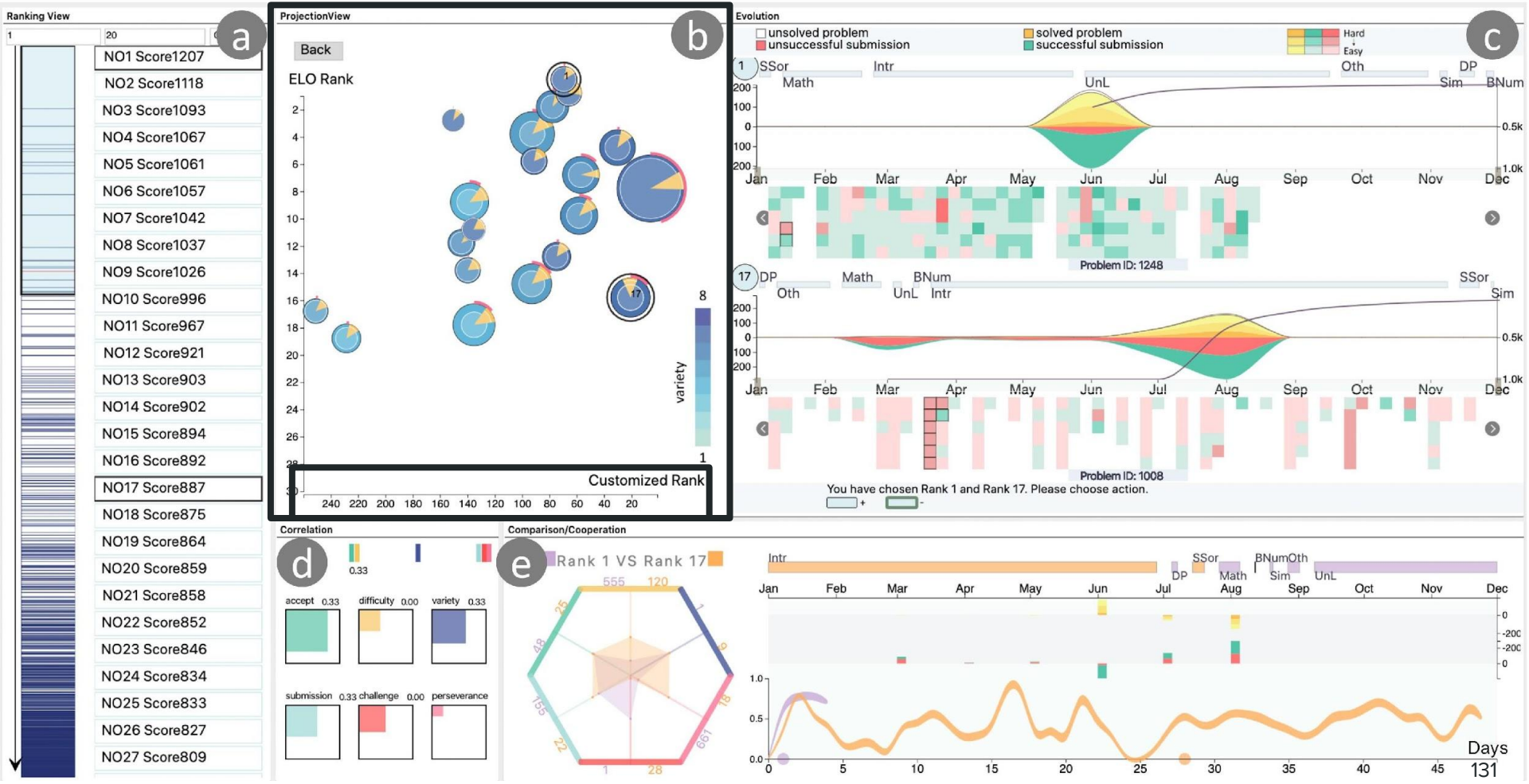

Ranking View
a
NO1 Score1207
NO2 Score1118
NO3 Score1093
NO4 Score1067
NO5 Score1061
NO6 Score1057
NO7 Score1042
NO8 Score1037
NO9 Score1026
ELO
NO13 Score903
NO14 Score902
NO15 Score894
NO16 Score892
NO17 Score887
NO18 Score875
NO19 Score864
NO20 Score859
NO21 Score858
NO22 Score852
NO23 Score846
NO24 Score834
NO25 Score833
NO26 Score827
NO27 Score809
ProjectionView
b
Back
ELO Rank
Customized Rank
variety
Evolution
c
unsolved problem
unsuccessful submission
solved problem
successful submission
Hard
Easy
Problem ID: 1248
Problem ID: 1008
You have chosen Rank 1 and Rank 17. Please choose action.
Correlation
d
accept
difficulty
variety
submission
challenge
perseverance
Comparison/Cooperation
e
Rank 1 VS Rank 17
Days
126

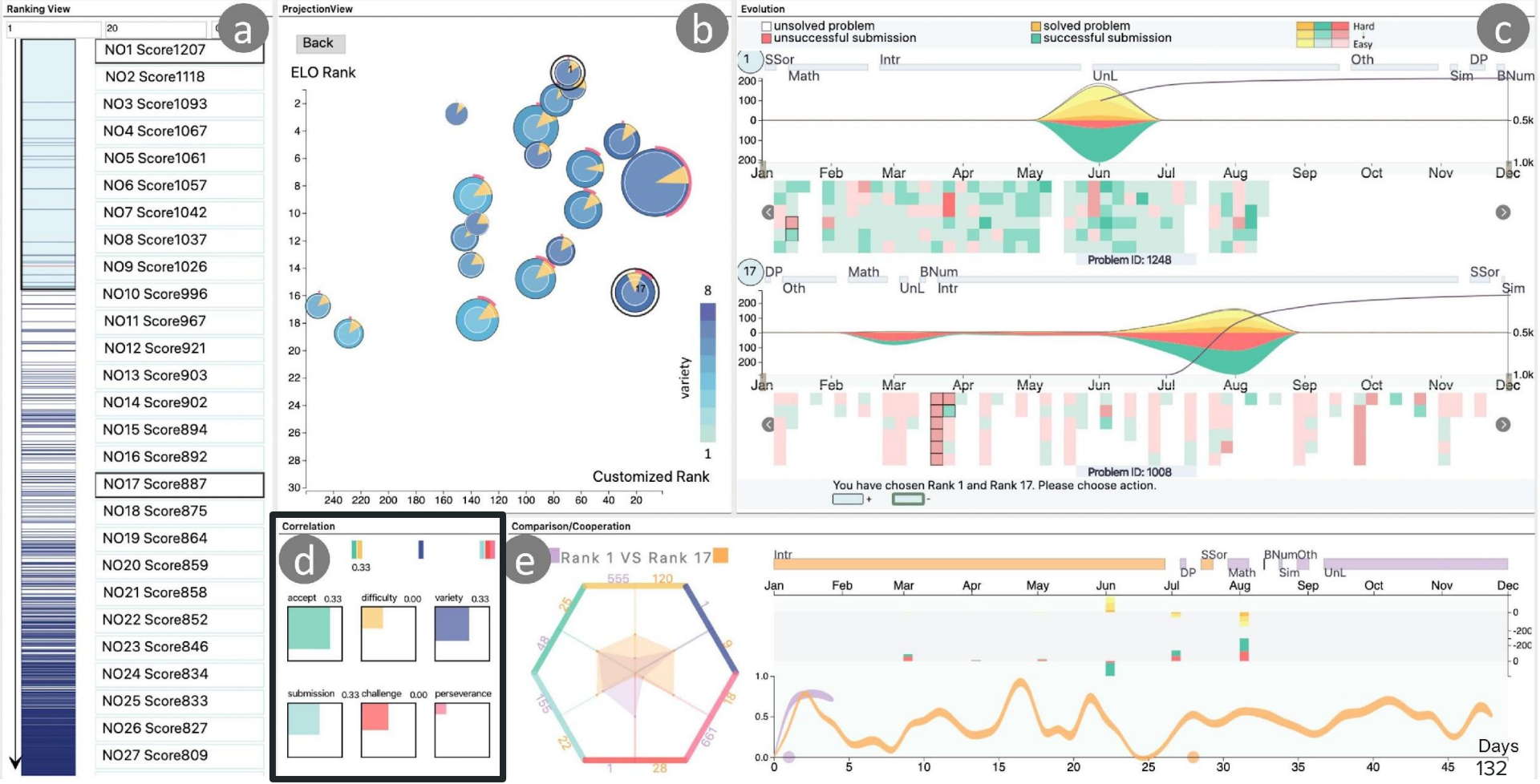

Ranking View
a
Rank 1
Rank 2
NO2 Score1118
NO4 Score1067
NO5 Score1061
NO6 Score1057
NO7 Score1042
NO8 Score1037
NO9 Score1026
NO10 Score996
NO11 Score967
NO12 Score921
NO13 Score903
NO14 Score902
NO15 Score894
NO16 Score892
NO17 Score887
NO18 Score875
NO19 Score864
NO20 Score859
NO21 Score858
NO22 Score852
NO23 Score846
NO24 Score834
NO25 Score833
NO26 Score827
NO27 Score809
ProjectionView
b
Back
ELO Rank
Customized Rank
variety
Evolution
c
unsolved problem
solved problem
unsuccessful submission
successful submission
Hard
Easy
Problem ID: 1248
Problem ID: 1008
You have chosen Rank 1 and Rank 17. Please choose action.
Correlation
d
accept
difficulty
variety
submission
challenge
perseverance
Comparison/Cooperation
e
Rank 1 VS Rank 17
Days
127

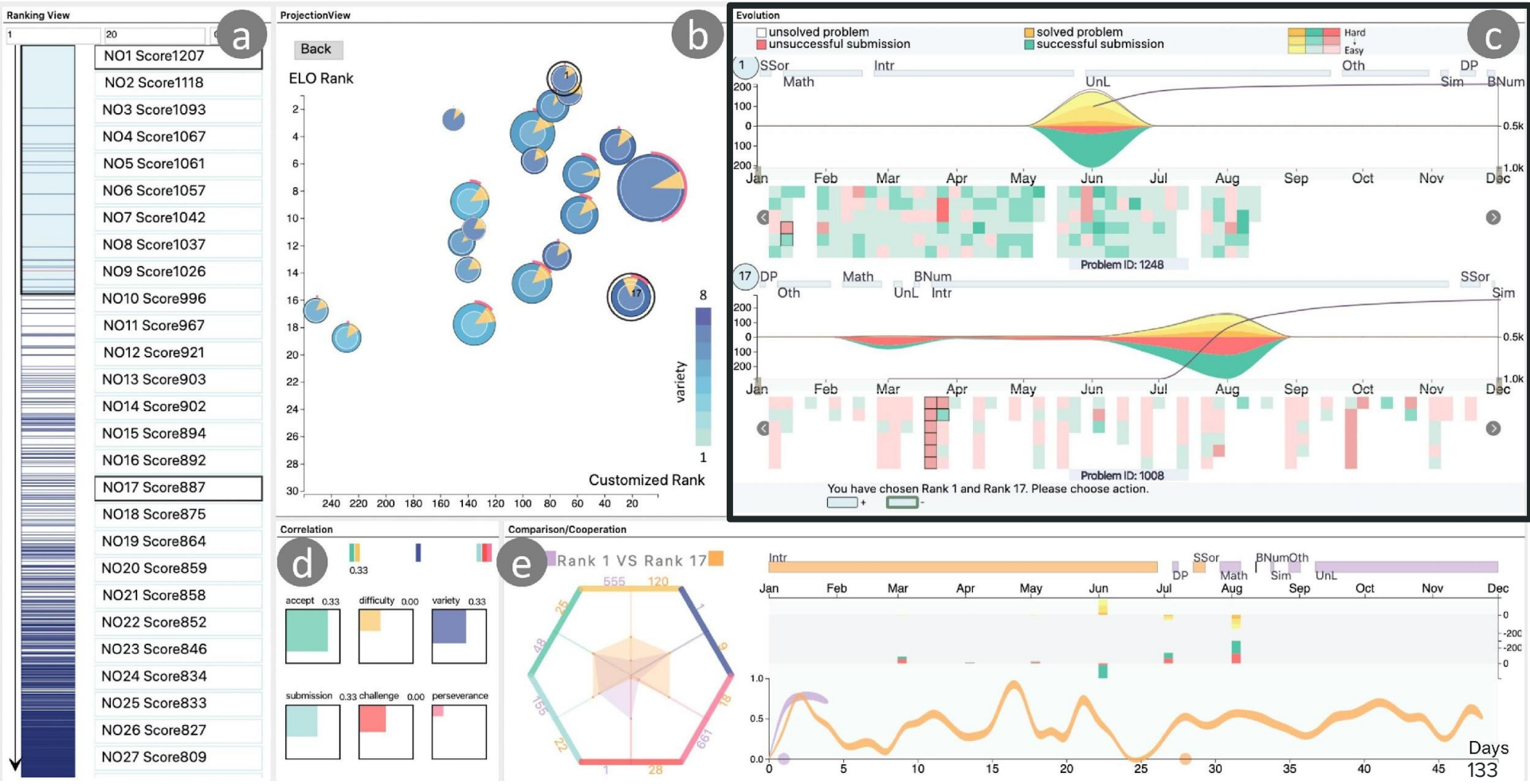

Ranking View
a
Rank 1
Rank 2
Rank 11
Rank 12
NO2 Score1118
NO4 Score1067
NO5 Score1061
NO6 Score1057
NO9 Score1026
NO10 Score996
NO11 Score967
NO12 Score921
NO13 Score903
NO14 Score902
NO15 Score894
NO16 Score892
NO17 Score887
NO18 Score875
NO19 Score864
NO20 Score859
NO21 Score858
NO22 Score852
NO23 Score846
NO24 Score834
NO25 Score833
NO26 Score827
NO27 Score809
ProjectionView
b
Back
ELO Rank
Customized Rank
variety
Evolution
c
unsolved problem
solved problem
unsuccessful submission
successful submission
Hard
Easy
Problem ID: 1248
Problem ID: 1008
You have chosen Rank 1 and Rank 17. Please choose action.
Correlation
d
accept
difficulty
variety
submission
challenge
perseverance
Comparison/Cooperation
e
Rank 1 VS Rank 17
Days
128

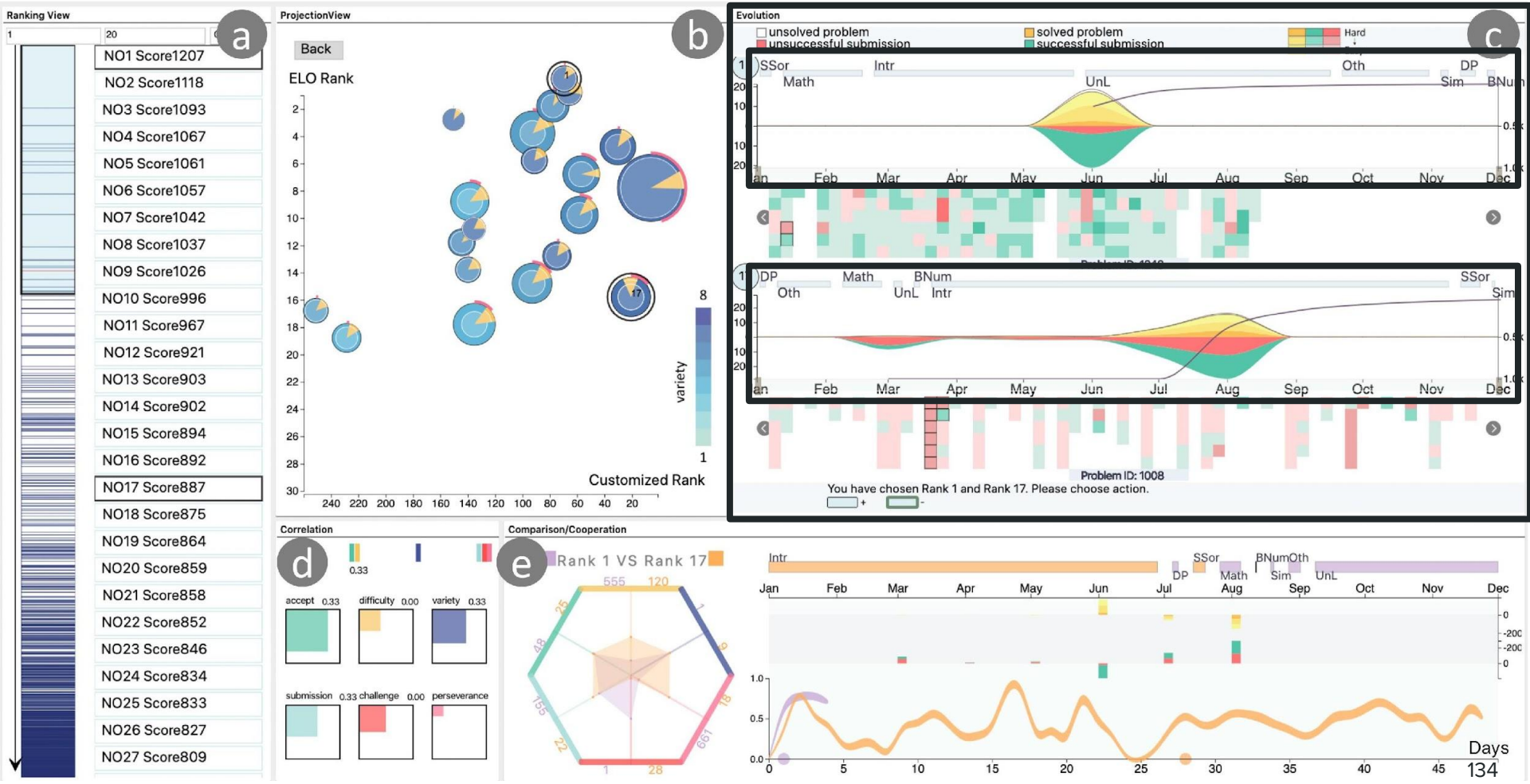

Ranking View
a
NO1 Score1207
NO2 Score1118
NO3 Score1093
NO4 Score1067
NO5 Score1061
NO6 Score1057
NO7 Score1042
NO8 Score1037
NO9 Score1026
NO10 Score996
NO11 Score967
NO12 Score921
NO13 Score903
NO14 Score902
NO15 Score894
NO16 Score892
NO17 Score887
NO18 Score875
NO19 Score864
NO20 Score859
NO21 Score858
NO22 Score852
NO23 Score846
NO24 Score834
NO25 Score833
NO26 Score827
NO27 Score809
ProjectionView
b
Back
ELO Rank
Customized Rank
variety
Evolution
c
unsolved problem
unsuccessful submission
solved problem
successful submission
Hard
Easy
Problem ID: 1248
Problem ID: 1008
You have chosen Rank 1 and Rank 17. Please choose action.
Correlation
d
accept
difficulty
variety
submission
challenge
perseverance
Comparison/Cooperation
e
Rank 1 VS Rank 17
Days
129

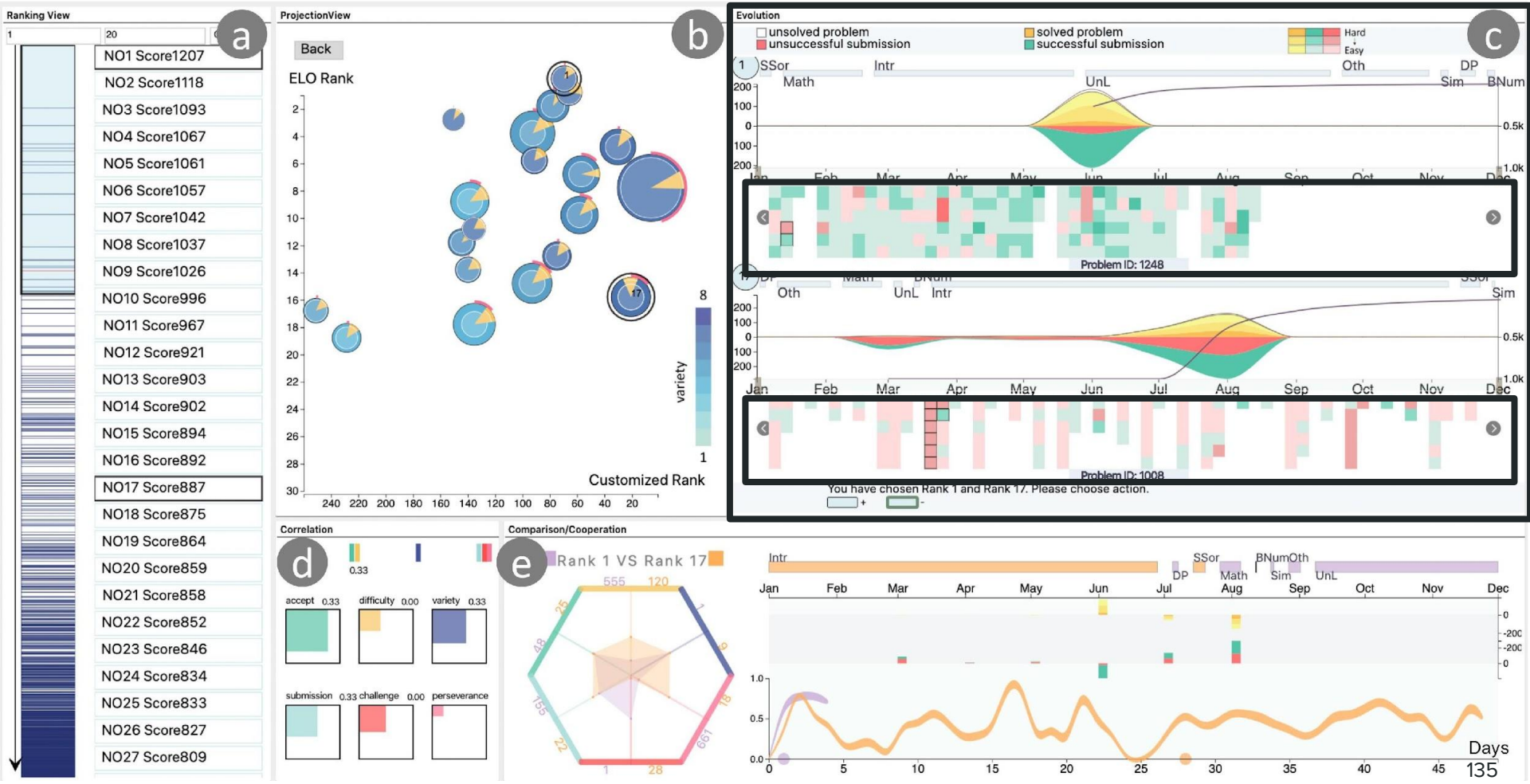

Ranking View
a
NO1 Score1207
NO2 Score1118
NO3 Score1093
NO4 Score1067
NO5 Score1061
NO6 Score1057
NO7 Score1042
NO8 Score1037
NO9 Score1026
NO10 Score996
NO11 Score967
NO12 Score921
NO13 Score903
NO14 Score902
NO15 Score894
NO16 Score892
NO17 Score887
NO18 Score875
NO19 Score864
NO20 Score859
NO21 Score858
NO22 Score852
NO23 Score846
NO24 Score834
NO25 Score833
NO26 Score827
NO27 Score809
ProjectionView
b
Back
ELO Rank
Customized Rank
variety
Evolution
c
unsolved problem
unsuccessful submission
solved problem
successful submission
Hard
Easy
SSor
Intr
Math
UnL
Oth
DP
Sim
BNum
Problem ID: 1248
Problem ID: 1008
You have chosen Rank 1 and Rank 17. Please choose action.
Correlation
d
accept
difficulty
variety
submission
challenge
perseverance
Comparison/Cooperation
e
Rank 1 VS Rank 17
Days
130

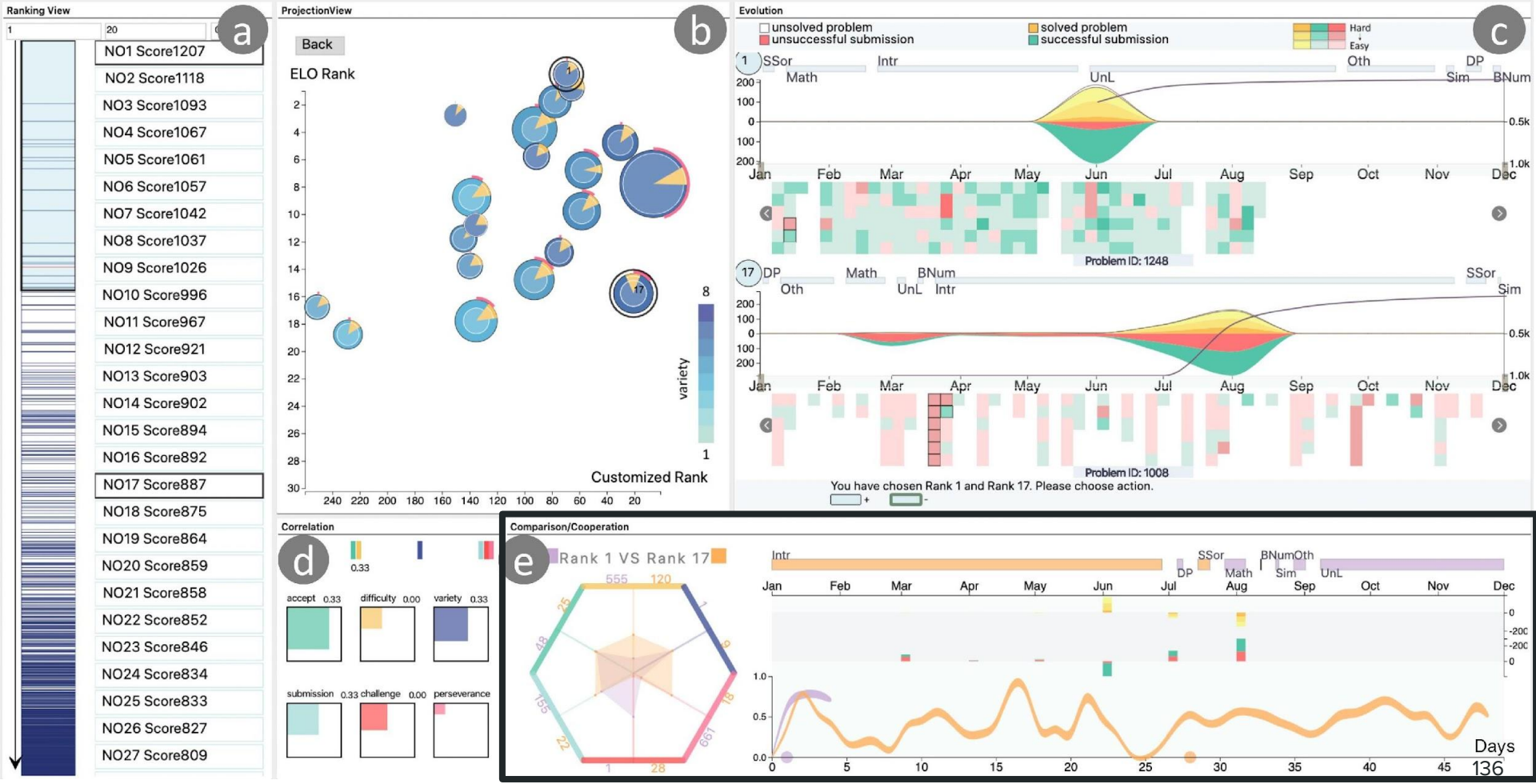

Ranking View
a
NO1 Score1207
NO2 Score1118
NO3 Score1093
NO4 Score1067
NO5 Score1061
NO6 Score1057
NO7 Score1042
NO8 Score1037
NO9 Score1026
NO10 Score996
NO11 Score967
NO12 Score921
NO13 Score903
NO14 Score902
NO15 Score894
NO16 Score892
NO17 Score887
NO18 Score875
NO19 Score864
NO20 Score859
NO21 Score858
NO22 Score852
NO23 Score846
NO24 Score834
NO25 Score833
NO26 Score827
NO27 Score809
ProjectionView
b
Back
ELO Rank
variety
Customized Rank
Evolution
c
unsolved problem
unsuccessful submission
solved problem
successful submission
Hard
Easy
SSor
Intr
Math
UnL
Oth
DP
Sim
BNum
Problem ID: 1248
Problem ID: 1008
You have chosen Rank 1 and Rank 17. Please choose action.
Correlation
d
accept
difficulty
variety
submission
challenge
perseverance
Comparison/Cooperation
e
Rank 1 VS Rank 17
Days
131

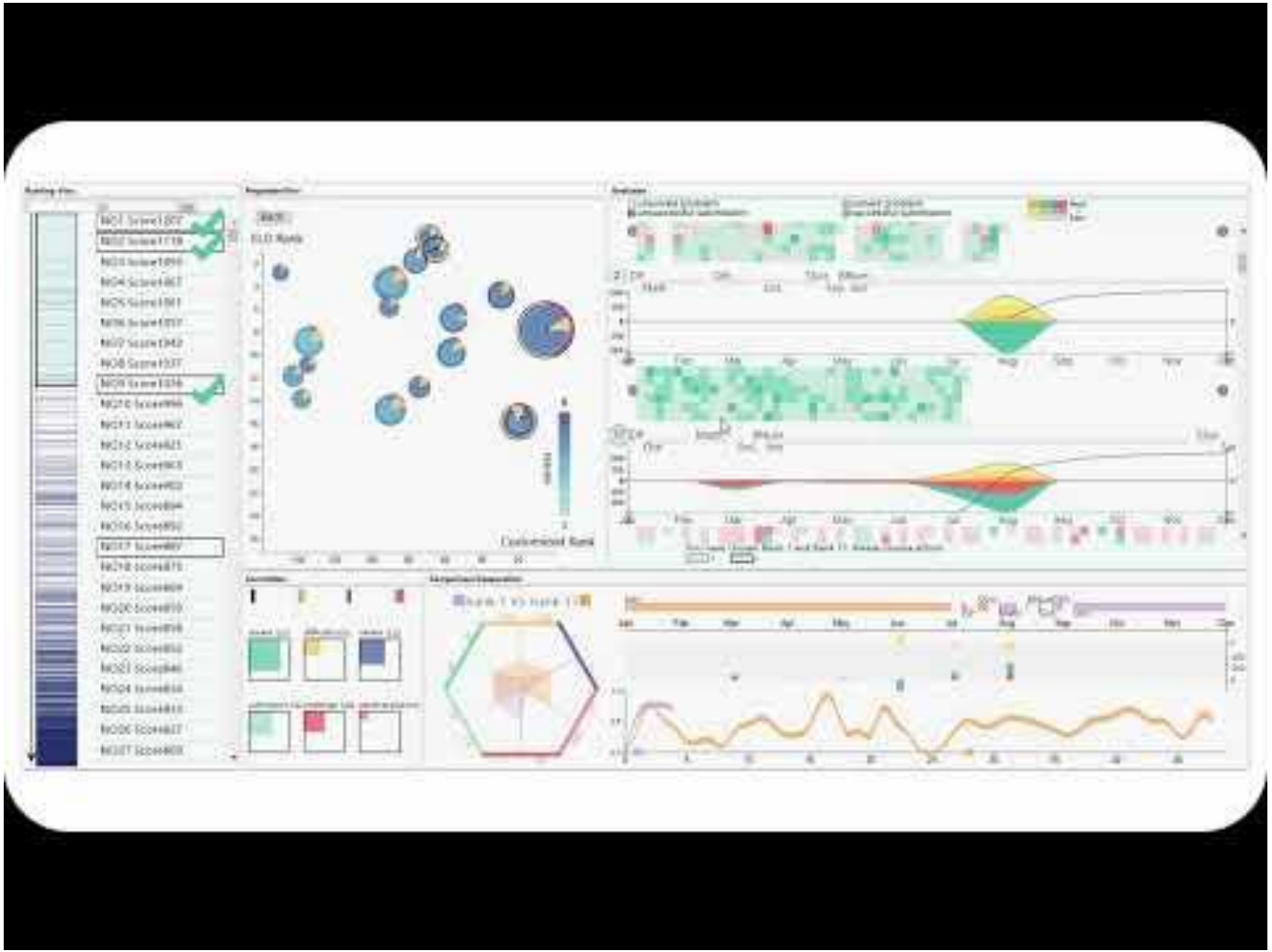

Ranking View
a
NO1 Score1207
NO2 Score1118
NO3 Score1093
NO4 Score1067
NO5 Score1061
NO6 Score1057
NO7 Score1042
NO8 Score1037
NO9 Score1026
NO10 Score996
NO11 Score967
NO12 Score921
NO13 Score903
NO14 Score902
NO15 Score894
NO16 Score892
NO17 Score887
NO18 Score875
NO19 Score864
NO20 Score859
NO21 Score858
NO22 Score852
NO23 Score846
NO24 Score834
NO25 Score833
NO26 Score827
NO27 Score809
ProjectionView
b
Back
ELO Rank
Customized Rank
variety
Evolution
c
unsolved problem
solved problem
unsuccessful submission
successful submission
Hard
Easy
SSor
Intr
Math
UnL
Oth
DP
Sim
BNum
Problem ID: 1248
Problem ID: 1008
You have chosen Rank 1 and Rank 17. Please choose action.
Correlation
d
accept 0.33
difficulty 0.00
variety 0.33
submission 0.33
challenge 0.00
perseverance
Comparison/Cooperation
e
Rank 1 VS Rank 17
Days
132


Ranking View
a
NO1 Score1207
NO2 Score1118
NO3 Score1093
NO4 Score1067
NO5 Score1061
NO6 Score1057
NO7 Score1042
NO8 Score1037
NO9 Score1026
NO10 Score996
NO11 Score967
NO12 Score921
NO13 Score903
NO14 Score902
NO15 Score894
NO16 Score892
NO17 Score887
NO18 Score875
NO19 Score864
NO20 Score859
NO21 Score858
NO22 Score852
NO23 Score846
NO24 Score834
NO25 Score833
NO26 Score827
NO27 Score809
ProjectionView
b
Back
ELO Rank
Customized Rank
variety
Evolution
c
unsolved problem
unsuccessful submission
solved problem
successful submission
Hard
Easy
Problem ID: 1248
Problem ID: 1008
You have chosen Rank 1 and Rank 17. Please choose action.
Correlation
d
accept
difficulty
variety
submission
challenge
perseverance
Comparison/Cooperation
e
Rank 1 VS Rank 17
Days
133


Ranking View
a
NO1 Score1207
NO2 Score1118
NO3 Score1093
NO4 Score1067
NO5 Score1061
NO6 Score1057
NO7 Score1042
NO8 Score1037
NO9 Score1026
NO10 Score996
NO11 Score967
NO12 Score921
NO13 Score903
NO14 Score902
NO15 Score894
NO16 Score892
NO17 Score887
NO18 Score875
NO19 Score864
NO20 Score859
NO21 Score858
NO22 Score852
NO23 Score846
NO24 Score834
NO25 Score833
NO26 Score827
NO27 Score809
ProjectionView
b
Back
ELO Rank
Customized Rank
variety
Evolution
c
unsolved problem
solved problem
Hard
Problem ID: 1008
You have chosen Rank 1 and Rank 17. Please choose action.
Correlation
d
accept 0.33
difficulty 0.00
variety 0.33
submission 0.33
challenge 0.00
perseverance
Comparison/Cooperation
e
Rank 1 VS Rank 17
Days
134


Ranking View
a
NO1 Score1207
NO2 Score1118
NO3 Score1093
NO4 Score1067
NO5 Score1061
NO6 Score1057
NO7 Score1042
NO8 Score1037
NO9 Score1026
NO10 Score996
NO11 Score967
NO12 Score921
NO13 Score903
NO14 Score902
NO15 Score894
NO16 Score892
NO17 Score887
NO18 Score875
NO19 Score864
NO20 Score859
NO21 Score858
NO22 Score852
NO23 Score846
NO24 Score834
NO25 Score833
NO26 Score827
NO27 Score809
ProjectionView
b
Back
ELO Rank
Customized Rank
variety
Evolution
c
unsolved problem
unsuccessful submission
solved problem
successful submission
Hard
Easy
Problem ID: 1248
Problem ID: 1008
You have chosen Rank 1 and Rank 17. Please choose action.
Correlation
d
accept
difficulty
variety
submission
challenge
perseverance
Comparison/Cooperation
e
Rank 1 VS Rank 17
Days
135


Ranking View
a
NO1 Score1207
NO2 Score1118
NO3 Score1093
NO4 Score1067
NO5 Score1061
NO6 Score1057
NO7 Score1042
NO8 Score1037
NO9 Score1026
NO10 Score996
NO11 Score967
NO12 Score921
NO13 Score903
NO14 Score902
NO15 Score894
NO16 Score892
NO17 Score887
NO18 Score875
NO19 Score864
NO20 Score859
NO21 Score858
NO22 Score852
NO23 Score846
NO24 Score834
NO25 Score833
NO26 Score827
NO27 Score809
ProjectionView
b
Back
ELO Rank
Customized Rank
variety
Evolution
c
unsolved problem
unsuccessful submission
solved problem
successful submission
Hard
Easy
Problem ID: 1248
Problem ID: 1008
You have chosen Rank 1 and Rank 17. Please choose action.
Correlation
d
accept
difficulty
variety
submission
challenge
perseverance
Comparison/Cooperation
e
Rank 1 VS Rank 17
Days
136



Evaluation

Three usage scenarios

Elite Analysis and Selection

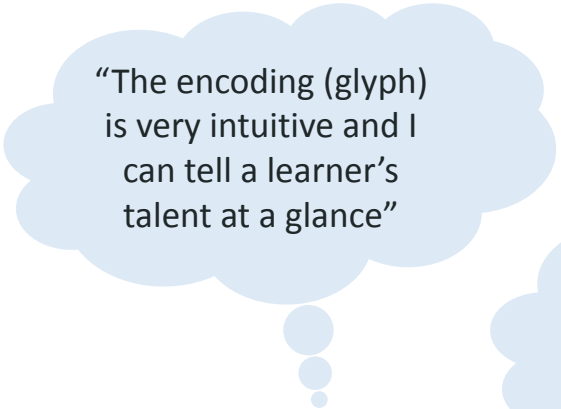
Personal Analysis and Training

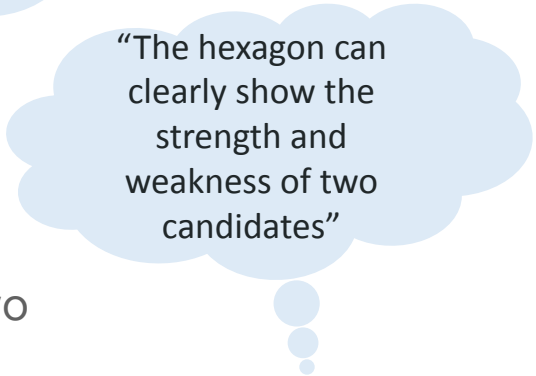
Team Formation

Five expert interviews

(Three coaches of competitive programming teams and two instructors teaching programming courses)

- ❑ System Usability
- ❑ System Effectiveness
- ❑ Visual Designs
- ❑ Interactions

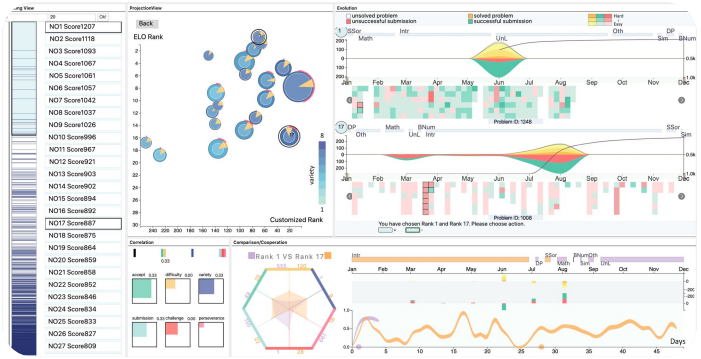




Overall, all five experts commented that SeqDynamics was **useful** and **easy to use**.

Conclusion

- An interactive visual analytical system to compare and rank objects with multiple temporal variables
- Novel glyphs and bilateral stacked graph for compsrison over different levels of detail
- Three usage scenarios and five expert interviews to show the system usefulness and usability



Summary

Learners (learning loop)

Educators (design loop)

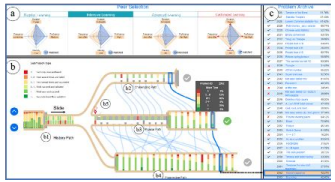
Customizing
learning goals and personalize activities

Reflecting
self-regulations on learning behaviors

Analyzing
learners behaviors and improve learning design

Data

- Large heterogeneous data
- Limited expertise and time in data analysis
- Not enough guidance and explanations
- No guarantee of data quality

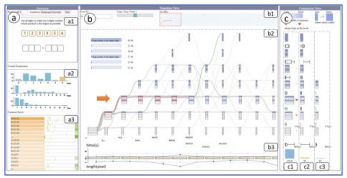


Peerlens (CHI 2019)

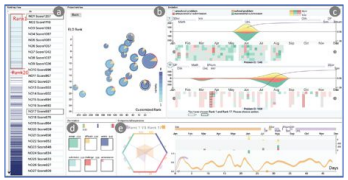
- Large heterogeneous data
- Limited expertise and time in data analysis
- Lack of motivation, consistent mental model, and actionable plans

RLens (L@S 2022)

- Large heterogeneous data
- Limited expertise and time in data analysis
- No predefined model



QLens (TVCG 2021)



SeqDynamics (EuroVIS 2020)

Future Work

Learners (learning loop)

Educators (design loop)

Customizing
learning goals and personalize activities

Reflecting
self-regulations on learning behaviors

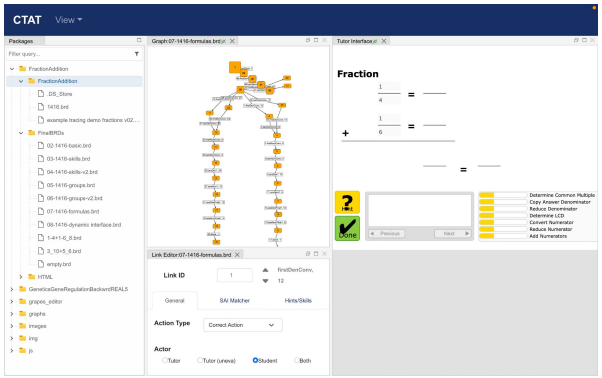
Analyzing
learners behaviors and improve learning design

Data

1. Integration of learning analytics and learning design

a. how to lower the barrier of learning design

b. how to support data-driven learning design



Future Work

Learners (learning loop)

Educators (design loop)

Customizing
learning goals and personalize activities

Reflecting
self-regulations on learning behaviors

Analyzing
learners behaviors and improve learning design

Data

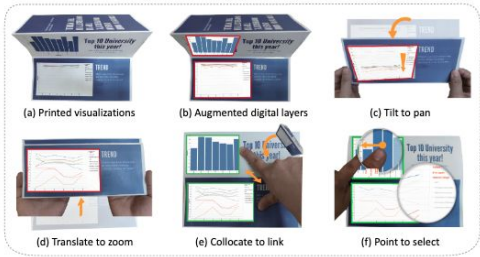
2. Real-time/synchronized personalized learning

a. Zoom: how to engage both instructors and learners?

b. Class-room/situated education: can we utilize immersive learning analytics?

Holstein, Kenneth, Bruce M. McLaren, and Vincent Aleven. "Student learning benefits of a mixed-reality teacher awareness tool in AI-enhanced classrooms." *International conference on artificial intelligence in education*. Springer, Cham, 2018.

Exploring Interactions with Printed Data Visualizations in Augmented Reality, (Tong et al., VIS 2022, conditionally accepted)



Future Work

Learners (learning loop)

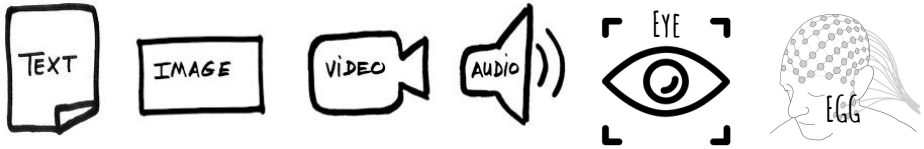
Educators (design loop)

Customizing
learning goals and
personalize activities

Reflecting
self-regulations on
learning behaviors

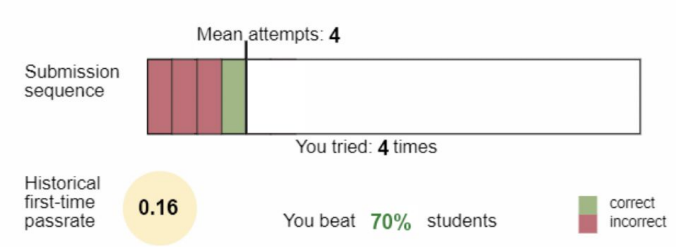
Analyzing
learners behaviors and
improve learning design

Data



3. Empower learners

a. how to infer learners' phycological state
b. how to motivate learners
c. how to design effective data visualization with low cognitive load
d. how to guarantee data quality [1]



[1] Choi, Kabdo, et al. "AlgoSolve: Supporting Subgoal Learning in Algorithmic Problem-Solving with Learnersourced Microtasks." *CHI Conference on Human Factors in Computing Systems*. 2022.

Xia, Meng, et al. "Using information visualization to promote students' reflection on" gaming the system" in online learning." *Proceedings of the Seventh ACM Conference on Learning@ Scale*. 2020.

Future Work

Learners (learning loop)

Educators (design loop)

Customizing
learning goals and
personalize activities

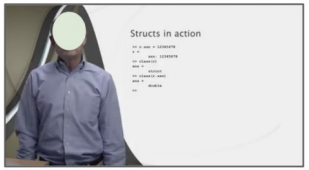
Reflecting
self-regulations on
learning behaviors

Data

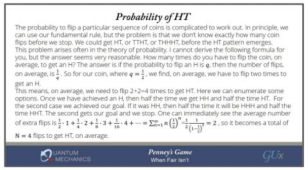
Analyzing
learners behaviors and
improve learning design

4. Personalization in diverse learning scenarios

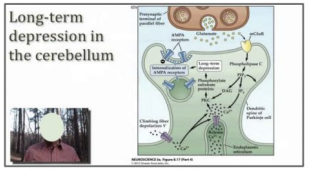
- ✓ Learning with different teachers or platforms [1]
- ✓ Learning with different hardware (smartphones, tablets, smart g
- × Learning with different scenaries (collaborative learning)



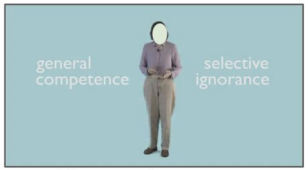
(a) Small font size



(b) Dense text



(c) Image containing text



(d) Low color contrast

[1] Xia, Meng, et al. "Understanding Distributed Tutorship in Online Language Tutoring." *LAK22: 12th International Learning Analytics and Knowledge Conference*. 2022.

Kim, Jeongyeon, et al. "Mobile-Friendly Content Design fo MOOCs: Challenges, Requirements, and Design Opportun *CHI Conference on Human Factors in Computing Systems* 2022. (***Best paper award***)

Future Work

Customizing
learning goals and personalize activities

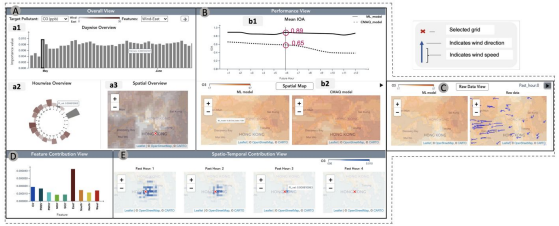
Reflecting
self-regulations on learning behaviors

Analyzing
learners behaviors and improve learning design

Data

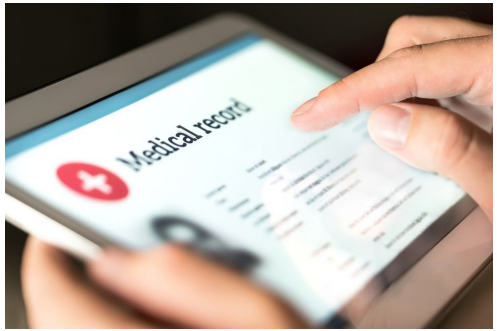
5. Visual analytics for personalization in other domains

XAI



AQX: Explaining Air Quality Forecast for Verifying Domain Knowledge using Feature Importance Visualization, Reshika et al., IUI 2022

Healthcare



Ongoing: Surgery Data Analysis

Fincance

Ongoing: NFT investiment strategy

Impact

[ ] DeepAI ...

Persua: A Visual Interactive System to Enhance the Persuasiveness of Arguments in Online Discussion

04/16/2022 · by Meng Xia, et al. ·
The Hong Kong University of Science and Technology ·
11 · share

Persuading people to change their opinions is a common practice in online discussion forums on topics ranging from political campaigns to relationship consultation. Enhancing people's ability to write persuasive arguments could not only practice their critical thinking and reasoning but also contribute to the effectiveness and civility in online communication. It is, however, not an easy task in online discussion settings where written words are the primary communication channel. In this paper, we derived four design goals for a tool

Using information visualization to promote students' reflection on" gaming the system" in online learning — 10 — 2020
M Xia, Y Asano, JJ Williams, H Qu, X Ma
Proceedings of the Seventh ACM Conference on Learning@ Scale, 37-49

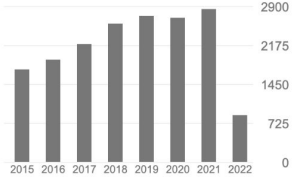
[HTML] Algorithmic bias in education
RS Baker, A Hawn - International Journal of Artificial Intelligence in …, 2021 - Springer
In this paper, we review algorithmic bias in education, discussing the causes of that bias and reviewing the empirical literature on the specific ways that algorithmic bias is known to have …
☆ Save Cite Cited by 22 Related articles All 2 versions

Ryan Baker — FOLLOW
University of Pennsylvania
Verified email at upenn.edu - Homepage
Educational Data Mining Learning Analytics Engagement Affect Learning

TITLE	CITED BY	YEAR
The state of educational data mining in 2009: A review and future visions RSJD Baker, K Yacef Journal of educational data mining 1 (1), 3-17	1906	2009
Learning analytics and educational data mining: towards communication and collaboration G Siemens, RSJ Baker Proceedings of the 2nd international conference on learning analytics and …	1180	2012
Educational data mining and learning analytics	1053	2014

Cited by — VIEW ALL

	All	Since 2017
Citations	22882	13972
h-index	68	54
i10-index	249	210



Publication List

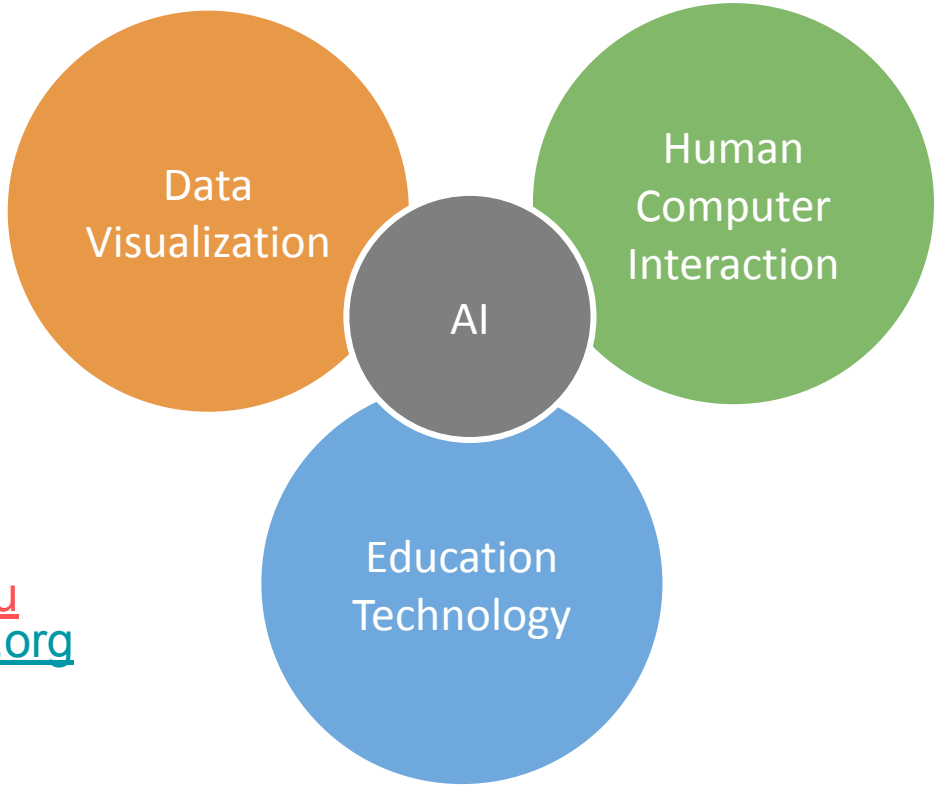
1. **Persua: A Visual Interactive System to Enhance the Persuasiveness of Arguments in Online Discussion**
   Meng Xia, Qian Zhu, Xingbo Wang, Fei Nie, Huamin Qu, Xiaojuan Ma, CSCW 2022
2. **RLens: A Computer-aided Visualization System for Supporting Reflection on Language Learning under Distributed Tutorship**
   Meng Xia, Yankun Zhao*, Jihyeong Hong*, Mehmet Hamza Erol*, Taewook Kim, Juho Kim, L@S 2022
3. **Understanding Distributed Tutorship in Online Language Tutoring**
   Meng Xia, Yankun Zhao, Mehmet Hamza Erol, Jihyeong Hong, Juho Kim, ACM LAK (Learning Analytics & Knowledge) 2022
4. **Exploring Interactions with Printed Data Visualizations in Augmented Reality**
   Wai Tong, Zhutian Chen, **Meng Xia**, Linping Yuan, Leo Yu Ho Lo, Benjamin Bach, Huamin Qu, VIS 2022 (conditionally accepted)
5. **Bias-Aware Design for Informed Decisions: Raising Awareness of Self-Selection Bias in User Ratings and Reviews**
   Qian Zhu, Leo Yu Ho Lo, **Meng Xia**, Zixin Chen, Xiaojuan Ma, CSCW 2022 (Accept with minor revision)
6. **Mobile-Friendly Content Design for MOOCs: Challenges, Requirements, and Design Opportunities**
   Jeongyeon Kim, Yubin Choi, **Meng Xia**, Juho Kim, CHI 2022, **Best Paper Award**
7. **“It Feels Like Taking a Gamble”: Exploring Perceptions, Practices, and Challenges of Using Makeup and Cosmetics for People with Visual Impairments**
   Mingzhe Li*, Franchesca Spector*, **Meng Xia***, Mina Oh*, Peter Cederberg, Yuqi Gong, Kristen Shinohara, Patrick Carrington, CHI 2022
8. **AlgoSolve: Supporting Subgoal Learning in Algorithmic Problem-Solving with Learnersourced Microtasks**
   Kabdo Choi, Hyungyu Shin, **Meng Xia**, Juho Kim, CHI 2022
9. **Explaining Air Quality Forecast for Verifying Domain Knowledge using Feature Importance Visualization**
   Reshika Palaniyappan Velumani, **Meng Xia**, Jun Han, Chaoli Wang, Alexis Lau, Huamin Qu, IUI 2022
10. **BlockLens: Visual Analytics of Student Coding Behaviors in Block-Based Programming Environments**
    Sean Tsung, Huan Wei, Haotian Li, **Meng Xia**, Yong Wang, Huamin Qu, L@S 2022 (Work In Progress)
11. **QLens: Visual Analytics of Multi-step Problem-solving Behaviors for Improving Question Design**
    Meng Xia, Reshika Palaniyappan Velumani, Yong Wang, Huamin Qu, Xiaojuan Ma, VIS 2020 (TVCG 2021)

Publication List

12. **Investigating the Effects of Robot Engagement Communication on Learning from Demonstration**
Mingfei Sun, Zhenhui Peng, **Meng Xia**, Xiaojuan Ma, International Journal of Social Robotics 2021
13. **Using Information Visualization to Promote Students' Reflection on "Gaming the system" in Online Learning**
Meng Xia, Yuya Asano, Joseph Jay Williams, Huamin Qu, Xiaojuan Ma, L@S 2020
14. **SeqDynamics: Visual Analytics for Evaluating Online Problem-solving Dynamics**
Meng Xia, Min Xu, Chuan-en Lin, Ta-ying Cheng, Huamin Qu, Xiaojuan Ma, EuroVIS 2020
15. **Predicting Student Performance in Interactive Online Question Pools Using Mouse Interactions**
Huan Wei, Haotian Li, **Meng Xia**, Yong Wang, Huamin Qu, ACM LAK (Learning Analytics & Knowledge) 2020
16. **Visual Analytics of Student Learning Behaviors on K-12 Mathematics E-learning Platforms**
Meng Xia, Huan Wei, Min Xu, Leo Yu Ho Lo, Yong Wang, Rong Zhang, Huamin Qu, IEEE VIS 2019 Poster, **Best Poster Award**
17. **PeerLens: Peer-inspired Interactive Learning Path Planning in Online Question Pool**
Meng Xia, Mingfei Sun, Huan Wei, Qing Chen, Yong Wang, Lei Shi, Huamin Qu, Xiaojuan Ma, CHI 2019
18. **Generation of Thangka Relief from Line Drawings**
Meng Xia, Rong Zhang, Ren Peng, Jinhui Yu, SCIENTIA SINICA Informationis 2018
19. **EnsembleLens: Ensemble-based Visual Exploration of Anomaly Detection Algorithms with Multidimensional Data**
Ke Xu, **Meng Xia**, Xing Mu, Yun Wang, Nan Cao, TVCG 2018
20. **Exploring How Software Developers Work with Mention Bot in GitHub**
Zhenhui Peng, Jeehoon Yoo, **Meng Xia**, Sunghun Kim, Xiaojuan Ma, in Proc. of Chinese CHI 2018
21. **Estimating Emotional Intensity from Body Poses for Human-Robot Interaction**
Mingfei Sun, Yiqing Mou, Hongwen Xie, **Meng Xia**, Michelle Wong, Xiaojuan Ma, in Proc. of IEEE SMC 2018
22. **Deep Spherical Panoramic Representation for 3D Shape Recognition**
Yuanli Feng, **Meng Xia**, Penglei Ji, Xiao Zhou, Ming Zeng, Xinguo Liu, Computer-Aided Design & Computer Graphics 2017
23. **Designing Kinect Game based on Video Tracking**
Yinglie Zhang, **Meng Xia**, Linqiang Chen, Computer Engineering and Applications 2015

Thank you!

Email: mengxia@andrew.cmu.edu
Homepage: https://www.xiameng.org



Projects

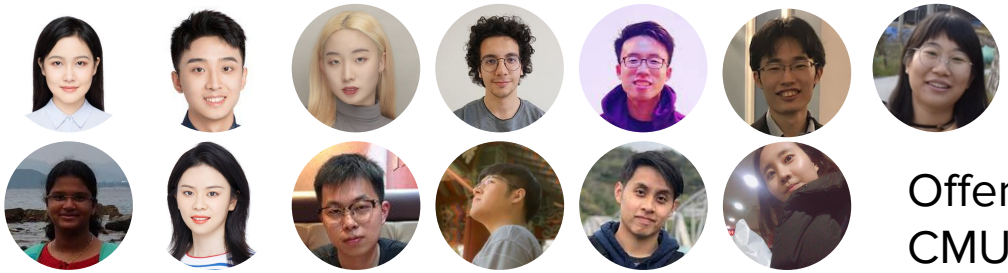
01	An Open Learning Design, Data Analytics and Visualization Framework for E-learning	• HKUST & HKU & MIT • Core Member • 2018 - 2021
02	Integration of Learning Design and Learning Analytics	• HKUST & HKU • Project Coordinator • 2020 - 2021
03	Anlaysis of Learning Progress and Recommendation of Personalized Learning Paths for English Learners	• KAIST • Project Coordinator • 2021 - 2022
04	Cognitive Tutor Authoring Tools (CTAT) for educational researchers	• Carnegine Mellon University • Project Coordinator • Since 2022

Advisors/
Mentors

Collaborators

7 countries

Interns /
Mentees



Offers from Stanford,
CMU, Oxford, etc.

Publication List

CCF A (9): CHI (4, including 1 **Best Paper**), TVCG (3), CSCW (2)

CORE A Conference in Education Technology (5) : LAK (2), L@S (3, including 1 Work in Progress)

CCF B (2) : IUI (1), EuroVis (1)

CCF C (1) : SMC (1)

CCF T1 (2) : SCIENTIA SINICA Informationis (1), CAD&CG (1)

Others (4): VIS (1 **Best Poster**), Computer Engineering and Applications (1), International Journal of Social Robotics (1), Chinese CHI (1)

Scholarships and Awards

KAIST

Best paper award at CHI 2022

HKUST

RGC Postdoctoral Fellowship (PDFS) 2021 (only 50 each year in HK)

SENG TOP RPg Award, 2018-2019

Best Poster Award at VIS, 2019

Overseas Research Award, 2018-2019

Zhejiang University

National Scholarship, 2015

Chairman of Postgraduate Association of Computer Science Department, 2014-

Outstanding graduate student and student cadres, 2014-2015

Hangzhou Dianzi University

National Scholarship, 2011

Best Poster Award at VIS, 2019

Best Paper Award at CHI, 2022

Teaching Plan

1. Data Visualization
2. Human Computer Interaction
3. Computer Organization
4. Computer Graphics
5. Personalized Online Learning

Professional Service

- Program Committee member for CHI 2023
- Program Committee member for VIS 2022
- Program Committee member for CHI 2022 LBW
- Program Committee member for VIS 2021

Patent List

1. Apparatus and Method for Evaluating Search Engine Performance, and Dashboard
   - KAIST, Jaehoon Lee, Juho Kim, Kabdo Choi, Mehmet Hamza Erol, Hyunwoo Kim, and **Meng Xia**, 10-2022-0026112
2. English conversation skill analysis using dialogue transcript
   - Jihyeong Hong, **Meng Xia**,Mehmet Hamza Erol, Juho Kim, KAIST, 10-2021-0106202
3. Utilizing tutor feedback for fine-grained learning progress reflection in online English tutoring via interactive visualization
   - **Meng Xia,** Jihyeong Hong, Mehmet Hamza Erol, Juho Kim, KAIST, 10-2021-0106212
4. QLens: Visual Analytics of Multi-step Problem-solving Behaviors for Improving Question Design
   - **Meng Xia,** Reshika Palaniyappan Velumani, Yong Wang, Huamin Qu, Xiaojuan Ma, Hong Kong University of Science and Technology, No.: US 63/102508
5. 一种将唐卡线描图生成浮雕效果的方法
   - 于金辉，**夏梦**，浙江大学, ZL 2015 1 1003097.2

Talks and presentations

CHI 2019

IEEE VIS 2019

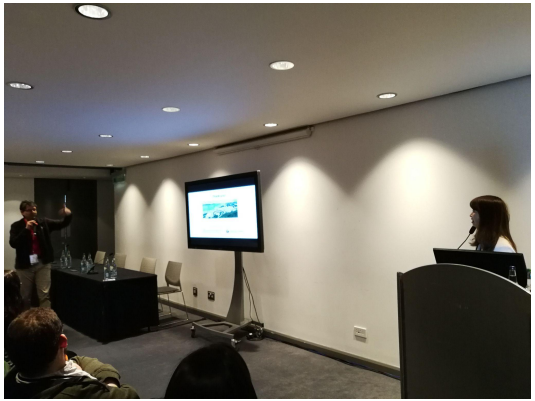
L@S 2020

Euro VIS 2020

IEEE VIS 2021

LAK 2022

L@S 2022



CHI 2019

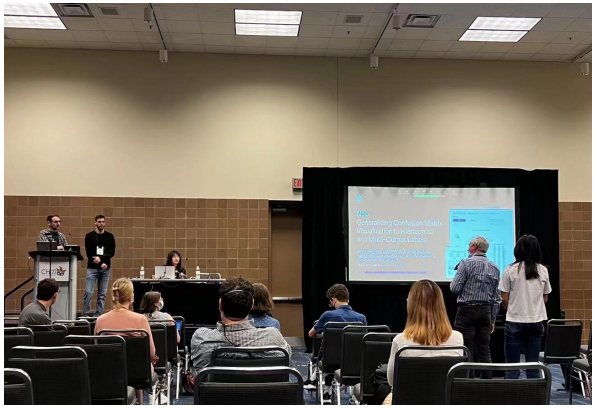
IEEE VIS 2019

CHI 2022: session chair

Invited Talk at KAIST HCI Course:

Visual Analytics and Its Application in Education

Invited Talk at VIS group at HKUST and ShanghaiTech University



IEEE CHI 2022

Plan on the Research Career

- Pushing forward the research in HCI and data visualization, particularly about personalized learning
- Build a team of undergraduate interns, master students, and PhD students from CS, education, and design
- Teaching courses about HCI, Data Visualization, Personalized Online Learning
- Apply for the Overseas Excellent Youth

Personalized Online Learning