

Human-centered Data-driven Systems for Personalized Online Learning

Meng Xia

06-01-2023



Online Learning has become widespread



A Coding Exercises
Platform

coursera

A MOOCs platform
(massive open online
courses)



A videotelephony
software



An online
discussion forum

Learning at scale, flexible locations and times

Scale



Machine Learning Specialization

#BreakIntoAI with Machine Learning Specialization. Master fundamental AI concepts and develop practical machine learning skills in the beginner-friendly, 3-course program by AI visionary Andrew Ng

★★★★★ 4.9 7,331 ratings

Andrew Ng [+3 more instructors](#) **TOP INSTRUCTORS**

120,420 already enrolled

Space



Time

CLASS TIMETABLE				
	9.00-9.30	9.30-10.30	11.00-1.00	1.30-2.30
MON	SPORT	MATHS	LITERATURE	ART
TUE	SPORT	LITERATURE	ENGLISH	HISTORY
WED	SPORT	SCIENCE	COMPUTING	MUSIC
THU	SPORT	ENGLISH	HISTORY	ART
FRI	SPORT	SCIENCE	COMPUTING	GEOGRAPHY



Learning at scale, flexible locations and times

Scale

Space

Time



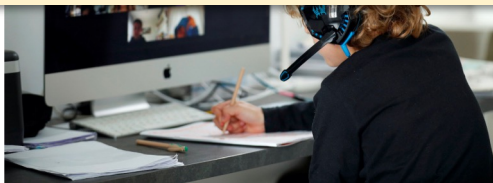
	9.00-9.30	9.30-10.30	11.00-1.00	1.30-2.30
MON	SPORT	MATHS	LITERATURE	ART
				HISTORY
				MUSIC
				ART
				GEOGRAPHY

Make learning tailored to individual students a grand challenge!

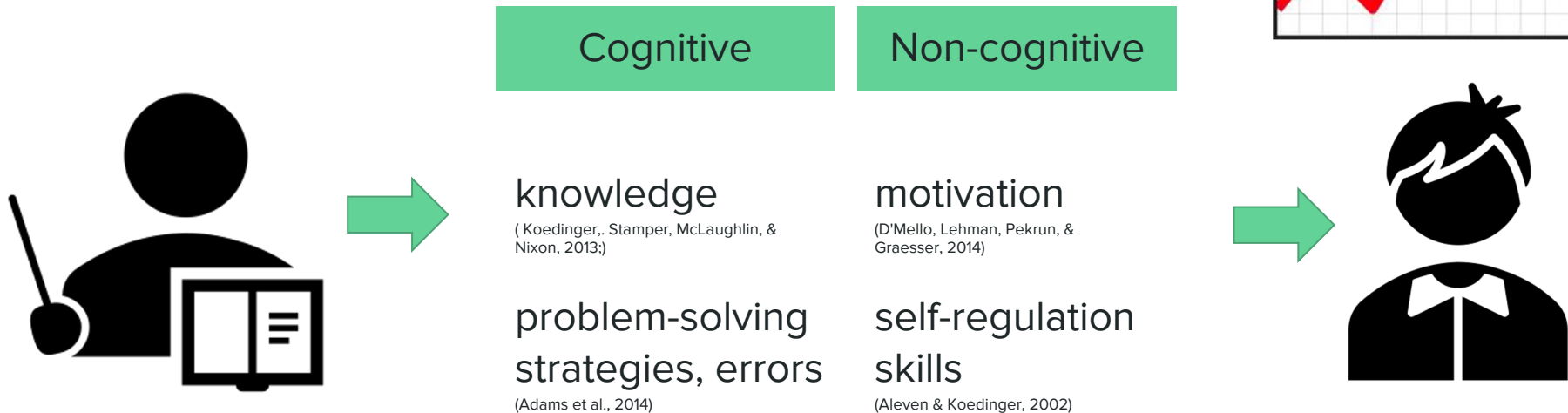
Machine Learning
#BreakIntoAI with Machine Learning
and develop practical machine learning
program by AI visionary Andrew Ng
★★★★★ 4.9 7,331 ratings

Andrew Ng +3 more instructors [TOP INSTRUCTORS](#)

120,420 already enrolled



Why personalized learning?



Learning at scale, flexible locations and times

Scale

Space

Time



CLASS TIMETABLE				
	9.00-9.30	9.30-10.30	11.00-1.00	1.30-2.30
MON	SPORT	MATHS	LITERATURE	ART
				HISTORY
				MUSIC
				ART
				GEOGRAPHY

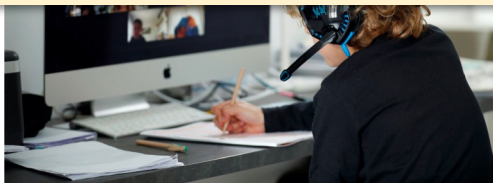
How could we achieve personalized online learning?

Machine Learning
#BreakIntoAI with Machine Learning
and develop practical machine learning
program by AI visionary Andrew Ng
★★★★★ 4.9 7,331 ratings

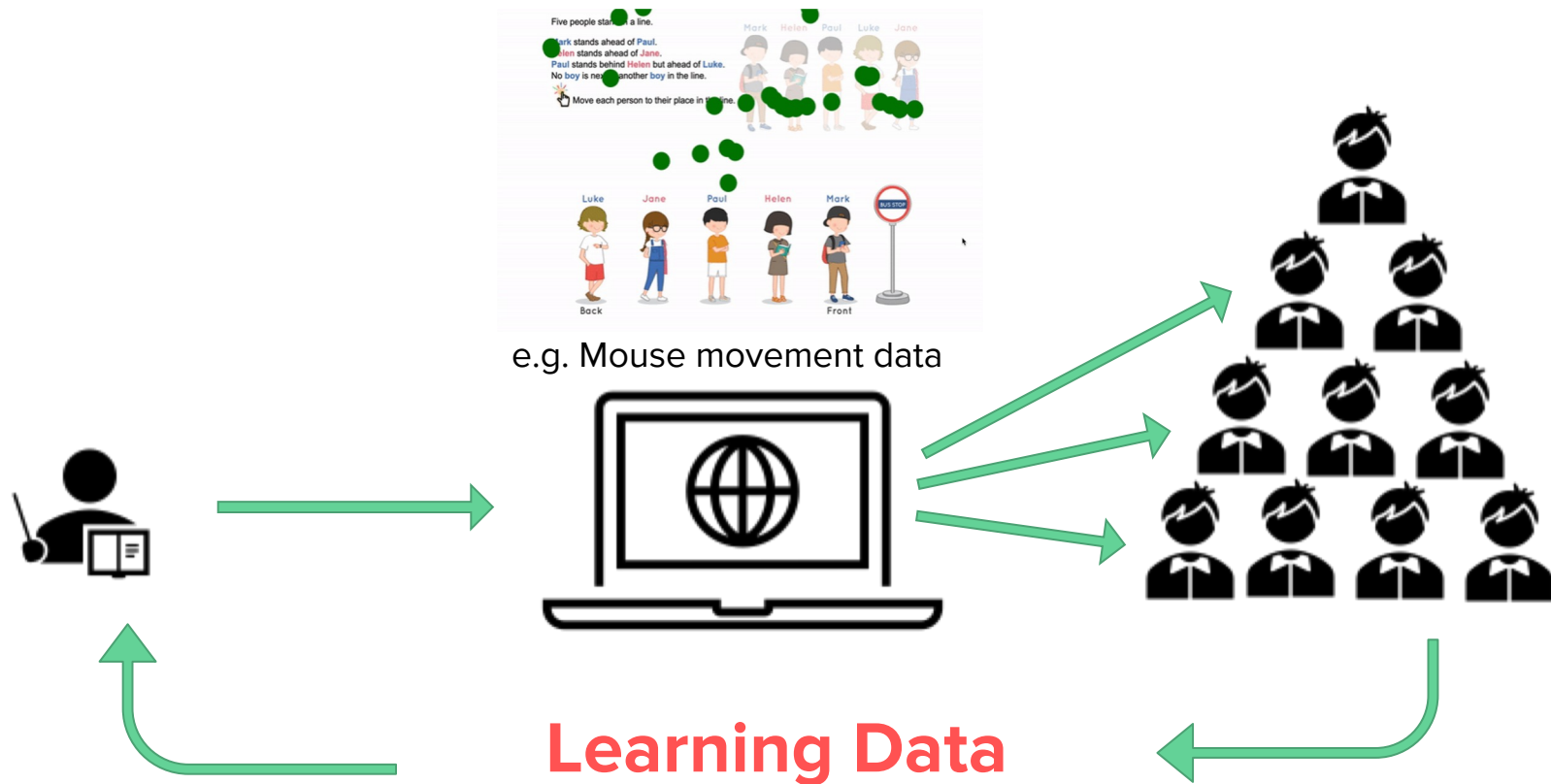


Andrew Ng +3 more instructors [TOP INSTRUCTORS](#)

120,420 already enrolled



Learning Data is a Communication Channel



Online Learning Platforms and Data



coursera

zoom



reddit



#	Title
✓ 1249	Minimum Remove to Make Valid Parentheses
✓ 1762	Buildings With an Ocean View
✓ 1570	Dot Product of Two Sparse Vectors
✓ 1650	Lowest Common Ancestor of a Binary Tree III
✓ 426	Convert Binary Search Tree to Sorted Doubly Linked List

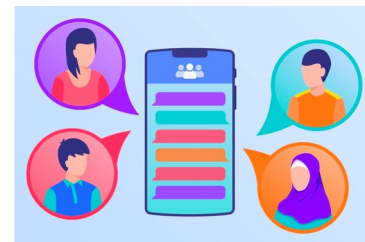
Problem-solving data



Click-stream data



Video and audio communication data



Online discussion data

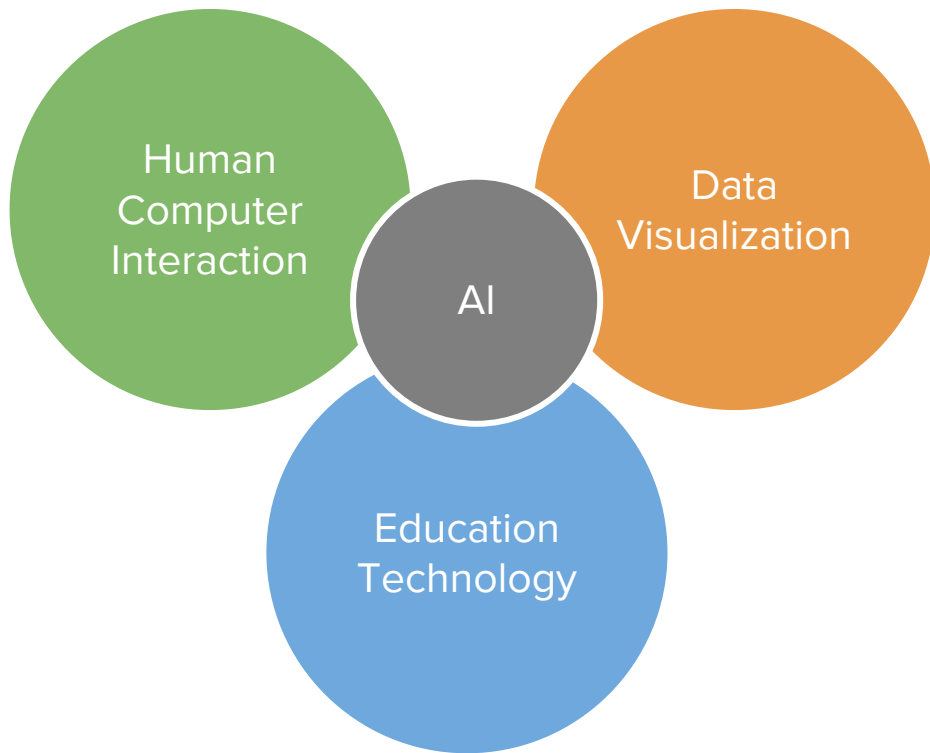
Research Goal

*Help **learners and educators** to **find insights** from learning data and use the insights to **make decisions** for achieving **personalized online learning**.*

Related Work on Learning Analytics

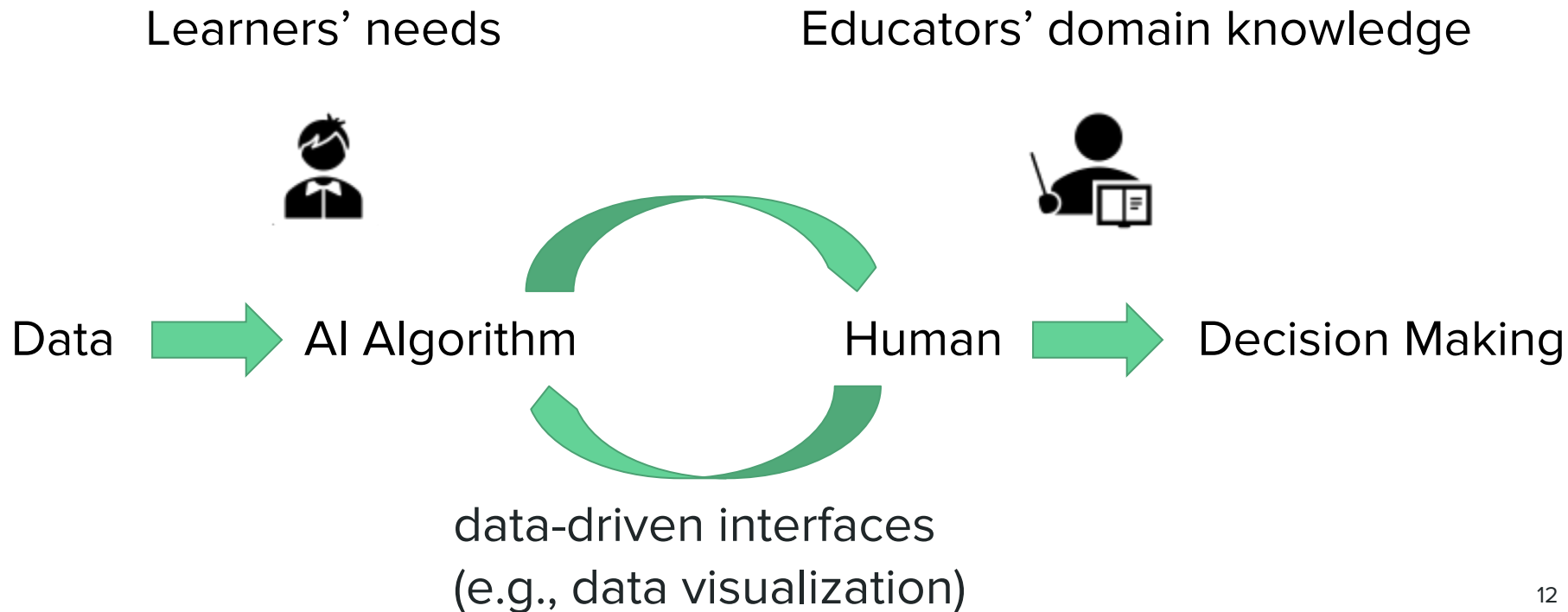
- Focus more on cognitive skills instead of non-cognitive variables (e.g., motivations, self-regulation skills)
Learning factor analysis (Cen et al., 2006) Performance factor analysis (Pavlik et al., 2009)
- Algorithms' outputs are not easy to interpret
Deep knowledge tracing (Piech et al., 2015) Explanatory models for educational data (Liu et al., 2017)

My Approach



*Design and build **human-centered data-driven systems** for achieving **personalized online learning**.*

Why *human-centered* data-driven systems?



Challenges

- **Learners have different aspects** that change over time
- **Learners and educators are not data analysts**
- Different learning scenarios:
 - Learning data is **voluminous and heterogeneous**
 - Learners and educators have **different tasks**



Learners (learning loop)

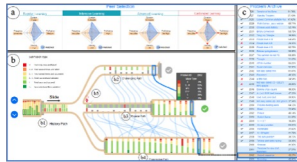
Planning

Performing Data

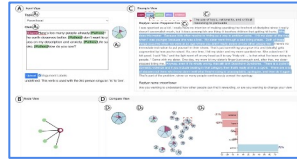
Reflecting

Understanding and Analyzing

Redesigning



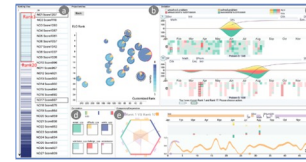
Peerlens (CHI 2019)



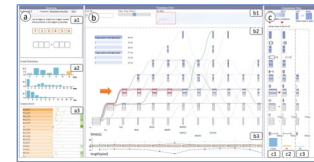
Persua (CSCW 2022)



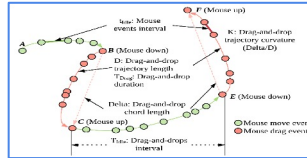
RLens (L@S 2022)



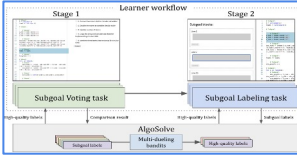
SeqDynamics (EuroVIS 2020)



QLens (TVCG 2021)



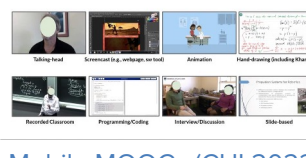
Predication (LAK 2020)



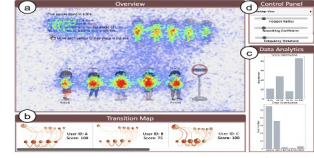
AlgoSolve (CHI 2022)



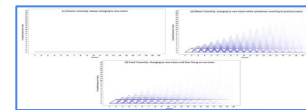
"Gaming the system" (L@S 2020)



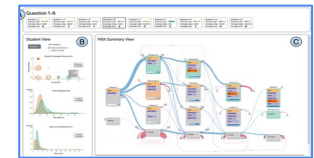
Mobile MOOCs (CHI 2022, Best Paper Award)



Visual Analytics K-12 (VIS 2019, Best Poster Award)



Distributed Tutorship (LAK 2022)



BlockLens (L@S 2022)

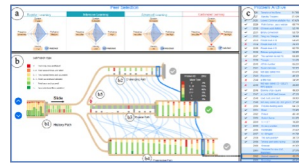


Learners (learning loop)

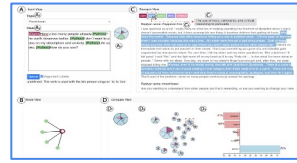
Planning

Performing
Data

Reflecting



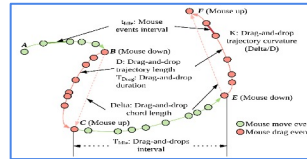
Peerlens (CHI 2019)



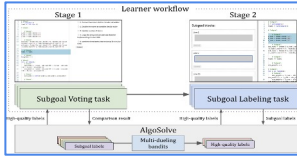
Persua (CSCW 2022)



RLens (L@S 2022)



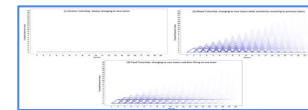
Predication (LAK 2020)



AlgoSolve (CHI 2022)



"Gaming the system" (L@S 2020)



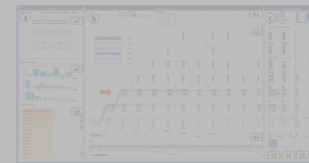
Distributed Tutorship (LAK 2022)

Understanding
and Analyzing

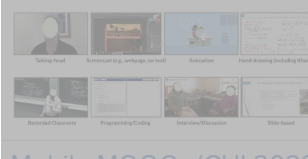
Redesigning



SeqDynamics (EuroVIS 2020)



QLens (TVCG 2021)



Mobile MOOCs (CHI 2022, Best Paper Award)



Visual Analytics K-12 (VIS 2019, Best Poster Award)



BlockLens (L@S 2022)



Educators (design loop)



Learners (learning loop)

Planning

Performing
Data

Reflecting

Understanding
and Analyzing

Redesigning



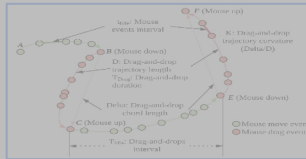
Peerlens (CHI 2019)



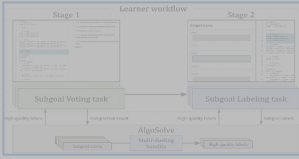
Persua (CSCW 2022)



RLens (L@S 2022)



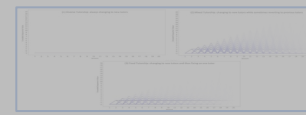
Predication (LAK 2020)



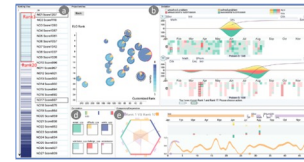
AlgoSolve (CHI 2022)



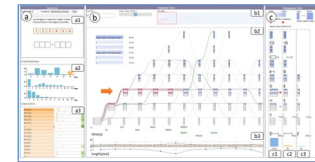
"Gaming the system" (L@S 2020)



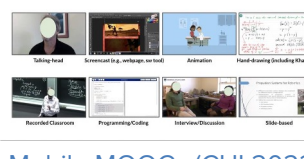
Distributed Tutorship (LAK 2020)



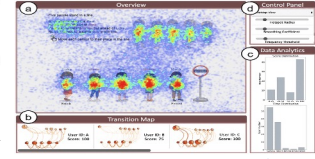
SeqDynamics (EuroVIS 2020)



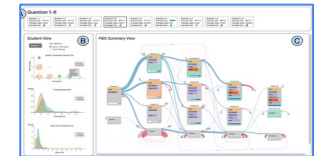
QLens (TVCG 2021)



Mobile MOOCs (CHI 2022, Best Paper Award)



Visual Analytics K-12 (VIS 2019, Best Poster Award)





Educators (design loop)



Learners (learning loop)

Planning

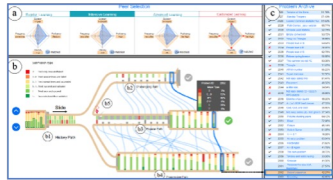
Performing

Data

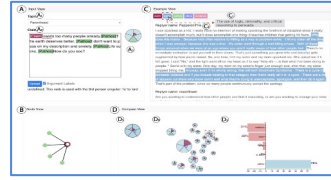
Reflecting

**Understanding
and Analyzing**

Redesigning



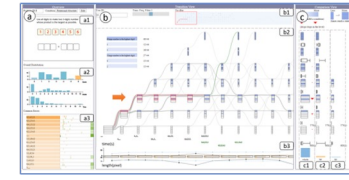
Peerlens (CHI 2019)



Persua (CSCW 2022)



RLens (L@S 2022)



QLens (TVCG 2021)

How data can be used for learners to plan, perform, and reflect on their learning?

How data can be used for educators to improve the design of learning materials?



Educators (design loop)



Learners (learning loop)

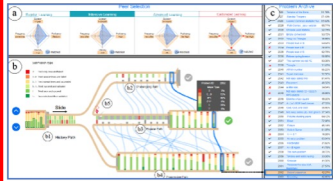
Planning

Performing
Data

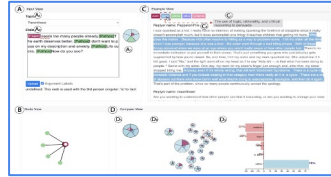
Reflecting

Understanding
and Analyzing

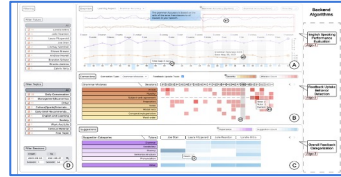
Redesigning



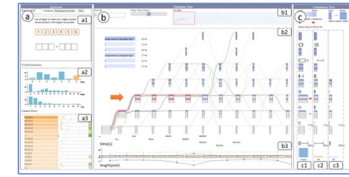
Peerlens (CHI 2019)



Persua (CSCW 2022)

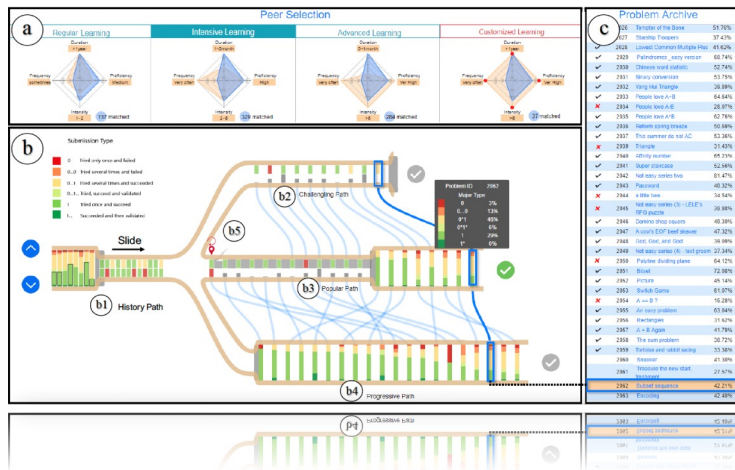


RLens (L@S 2022)



QLens (TVCG 2021)

How data can be used for learners to
plan, perform, and reflect on their
learning?



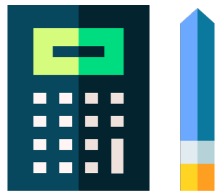
PeerLens: Peer-inspired Interactive Learning Path Planning in Online Question Pool

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

CHI 2019

What is an online question pool?

- A collection of questions for learners to practice their knowledge online



Math



Programming



Driving license

Motivation

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

Questions Pools:

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

Learners:

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

What to do next? What sequence to follow?

Related work: Educational Recommendation Techniques

(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)

=> **We lack the problem label information for use**

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)

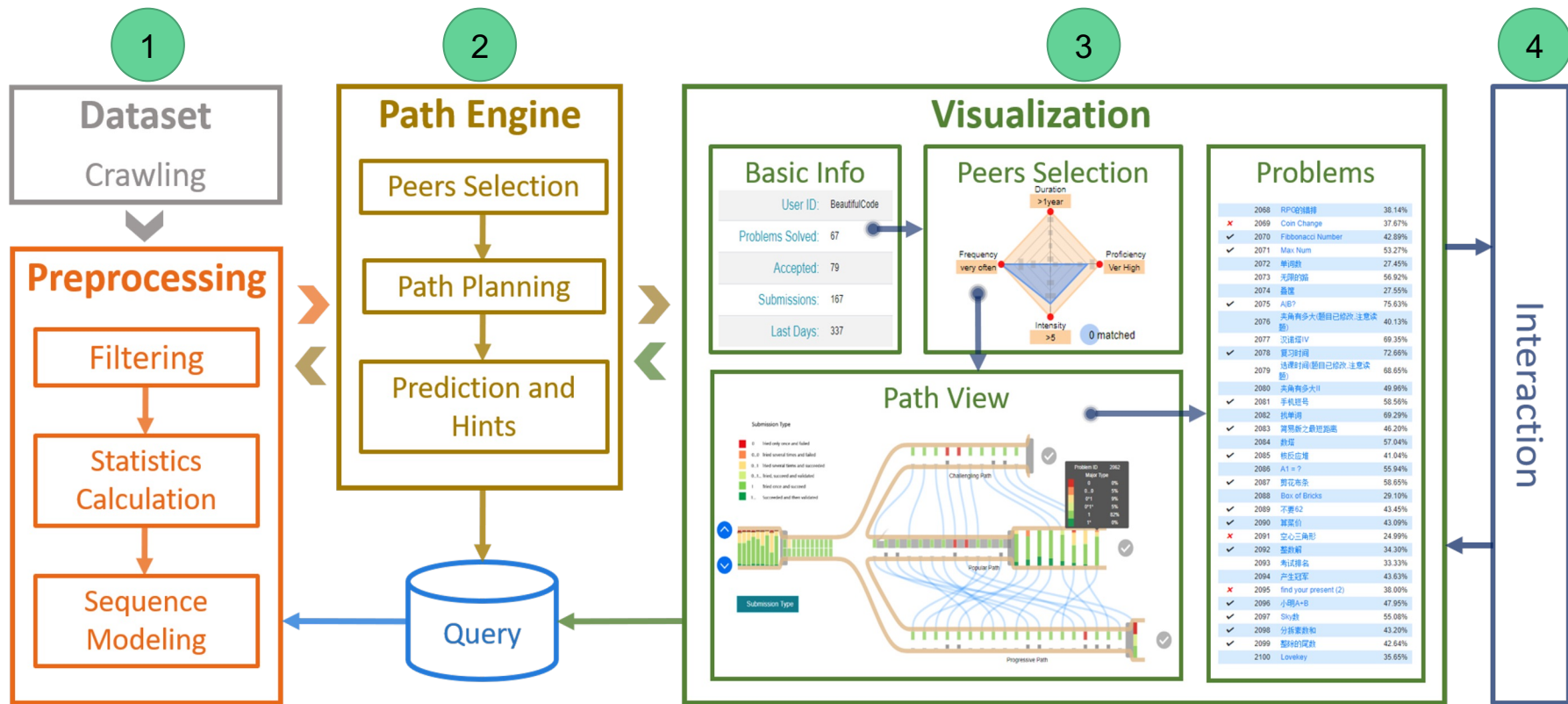
=> **Learners' intent is not well considered, and no explanation is provided for the results**

How learning data can be used for planning the learning path?

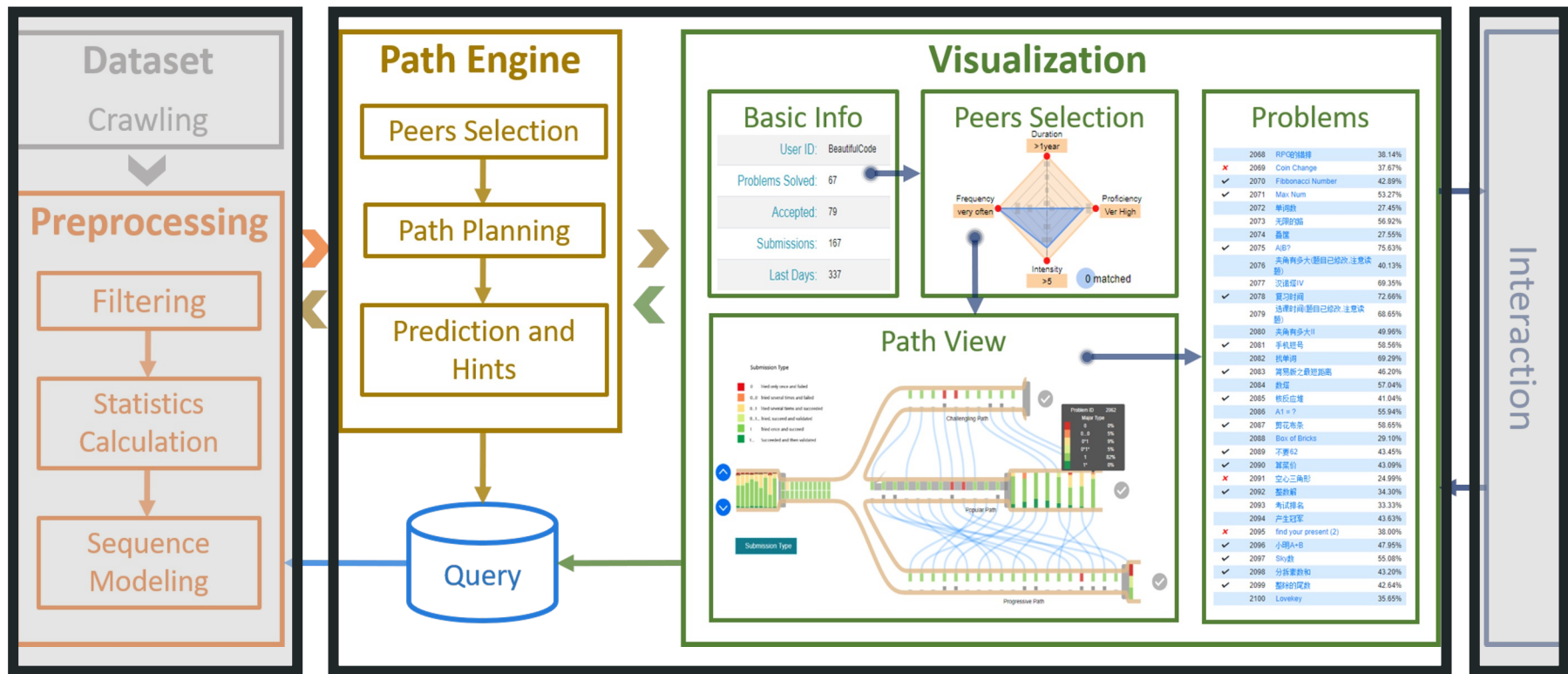
A user-centered design process

- Four participants: two question pool designers and two users
- Requirements gathering iteratively for three months
 - R1: **Find peers** for a target learning scenario.
 - R2: **Compare with peers'** performance to understand the gap.
 - R3: Offer flexible learning **path suggestions with explanations.**
 - R4: Provide **convenient interaction and intuitive visual designs.**

PeerLens System overflow



System overflow



How to quantify learner's performance?

R1: **Find peers** for a target learning scenario.

R2: **Compare with peers'** performance to understand the gap.

Difficulty of the questions

Accept rate?



Q1



1/5 => Hard problem

Q1

P1



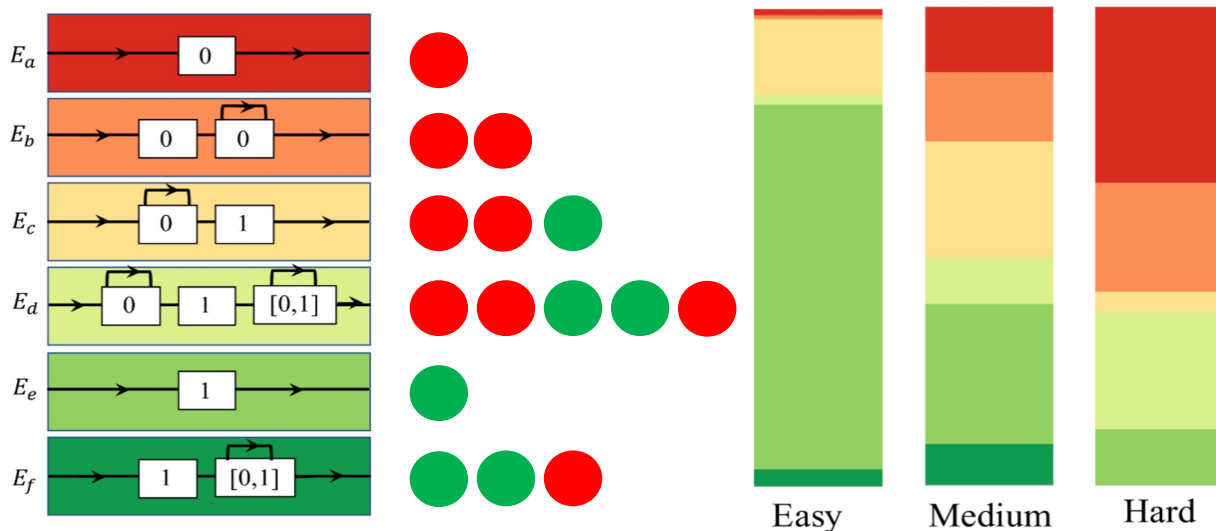
P2



Accept rate: 1/5 => **Still a hard problem?**

How to quantify learner's performance?

Submission type: the way a learner submits a question. ✓



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

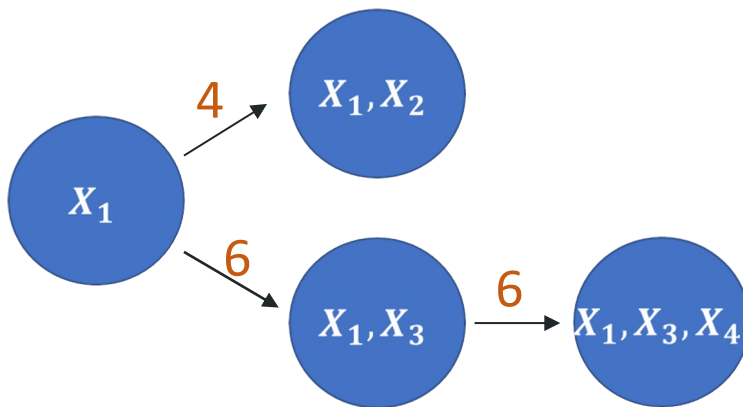
incorrect correct

How to recommend?

R1: **Find peers** for a target learning scenario.

R2: **Compare with peers'** performance to understand the gap.

R3: Offer flexible learning **path suggestions** with explanations.



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

How to recommend?

R1: **Find peers** for a target learning scenario.

R2: **Compare with peers'** performance to understand the gap.

R3: Offer flexible learning **path suggestions** with explanations.

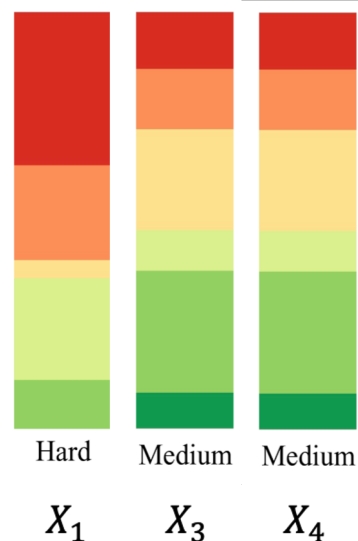
Popular path: $X_1 \rightarrow X_3 \rightarrow X_4$



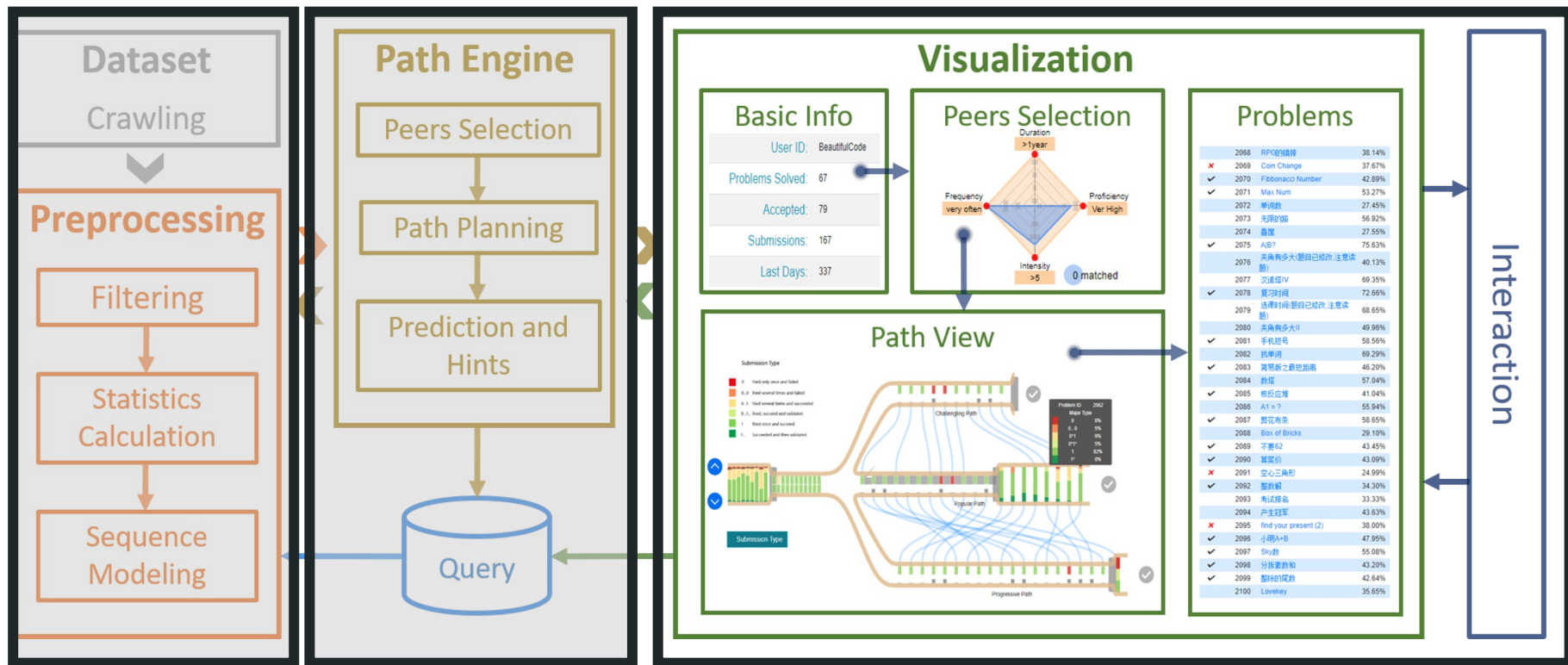
Challenging path: $X_1 \rightarrow X_4$



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



System overflow



Basic Info

User ID:

Problems Solved: 67

Accepted: 79

Submissions: 167

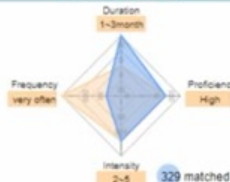
Last Days: 337

Peer Selection

Regular Learning



Intensive Learning



Advanced Learning



Customized Learning

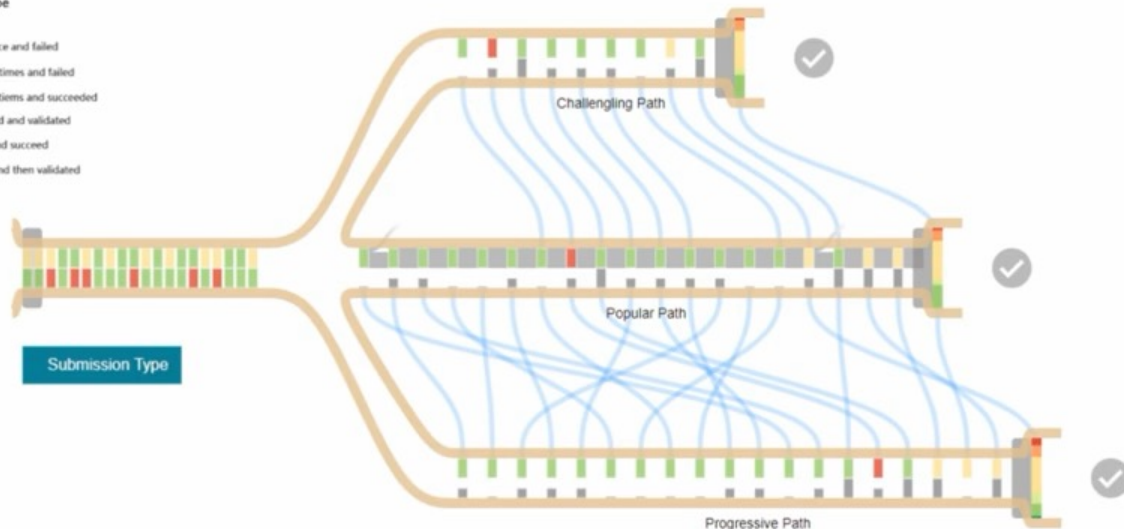


Problem Archive

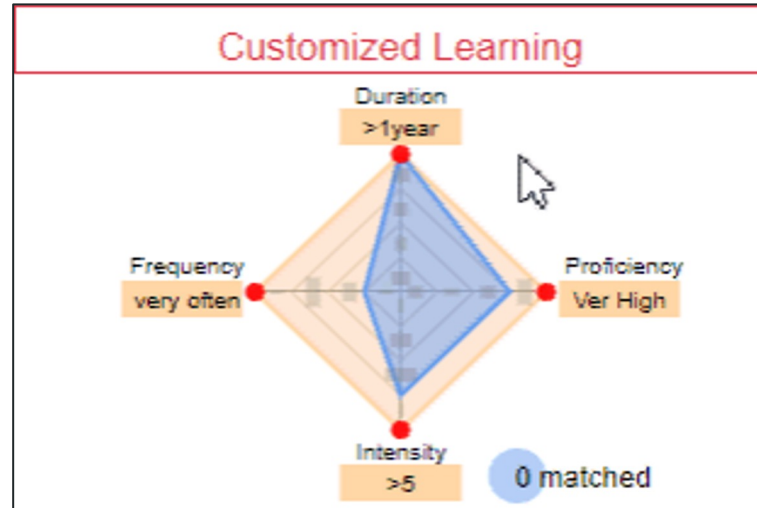
✓	2010	水仙花数	27.56%
✓	2011	多项式求和	56.00%
✓	2012	素数判定	36.26%
✓	2013	蜂猴记	74.48%
✗	2014	青年歌手大奖赛_评委打分	48.51%
✓	2015	偶数求和	39.52%
✗	2016	数据的交换输出	36.36%
✓	2017	字符串统计	55.48%
✓	2018	奶牛的故事	48.10%
✓	2019	数列有序I	40.76%
✗	2020	绝对值排序	45.44%
✓	2021	发工资咯：)	50.96%
✓	2022	海选女主角	41.82%
✗	2023	求平均成绩	21.66%
✓	2024	C语言合法标识符	36.85%
✓	2025	查找最大元素	51.91%
✗	2026	数字字母大小写	51.76%
✓	2027	统计元素	37.43%
✓	2028	Lowest Common Multiple Plus	41.62%
✓	2029	Palindromes_easy version	60.74%
✗	2030	汉字统计	52.74%
✓	2031	进制转换	53.75%
✓	2032	杨辉三角	39.89%
✓	2033	人见人爱A+B	64.64%
✗	2034	人见人爱A*B	28.07%
✓	2035	人见人爱A^B	62.76%
✓	2036	改革春风吹满地	50.69%
✗	2037	今年暑假不AC	53.36%
✓	2039	三角形	31.43%
✓	2040	景和数	65.23%
✗	2041	超级楼梯	52.56%
✓	2042	不容易系列之二	81.47%
✓	2043	密码	40.32%
✓	2044	一只小蜜蜂...	34.54%
✗	2045	不容易系列之(3)——LELE的RPG难题	39.98%

Submission Type

- 0 tried only once and failed
- 0.0 tried several times and failed
- 0.1 tried several times and succeeded
- 0.1... tried, succeed and validated
- 1 tried once and succeed
- 1... Succeeded and then validated



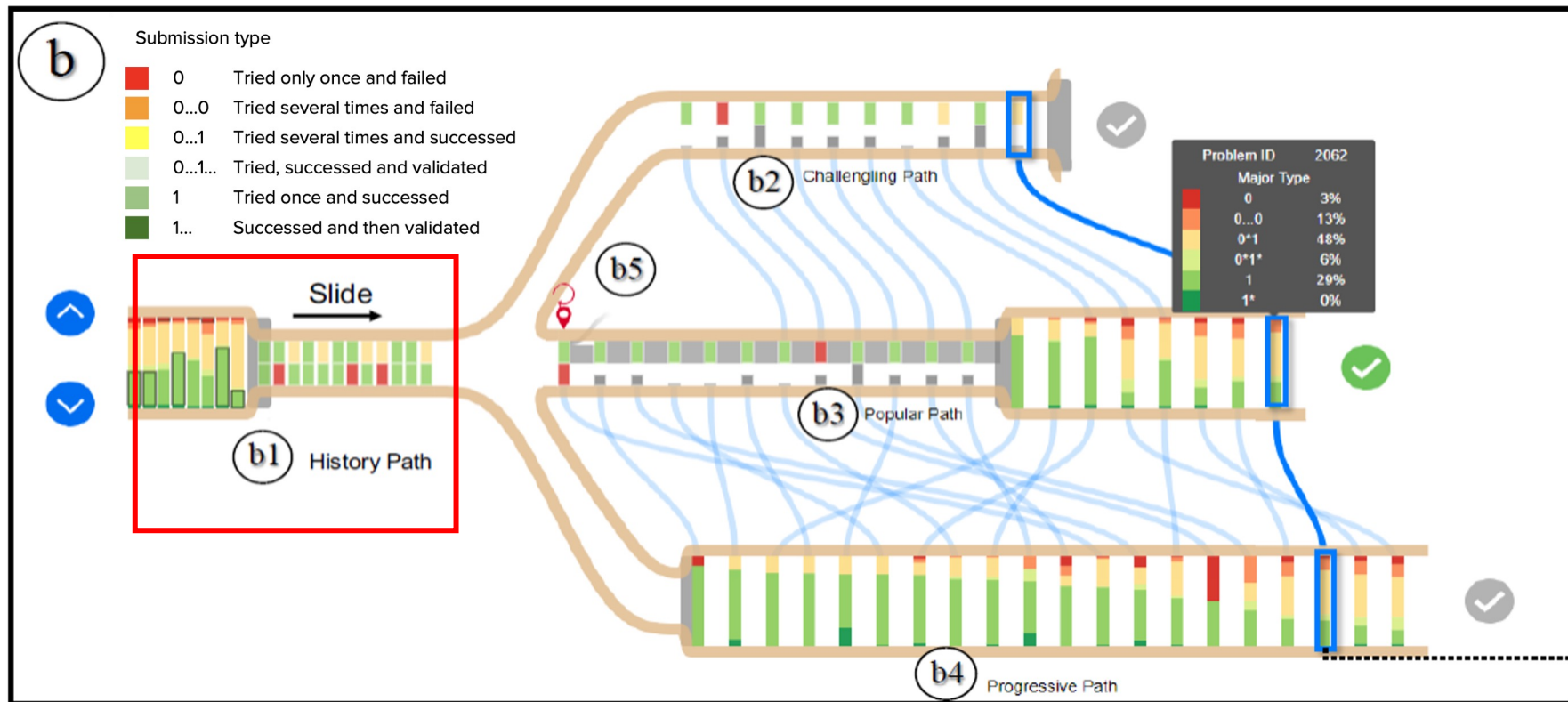
Visual Design: Peer Selection View



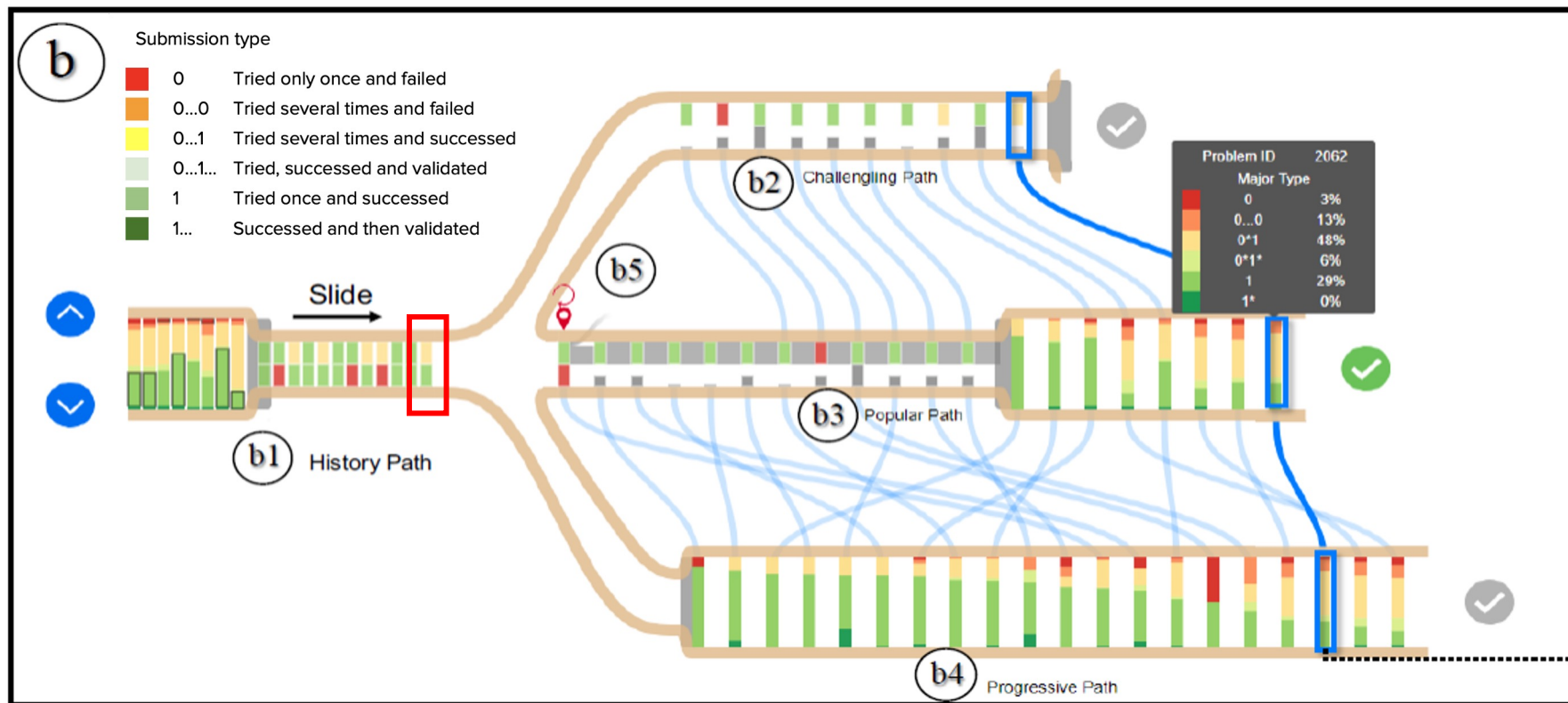
Yellow diamond plot: selected peers

Blue diamond plot: learner himself

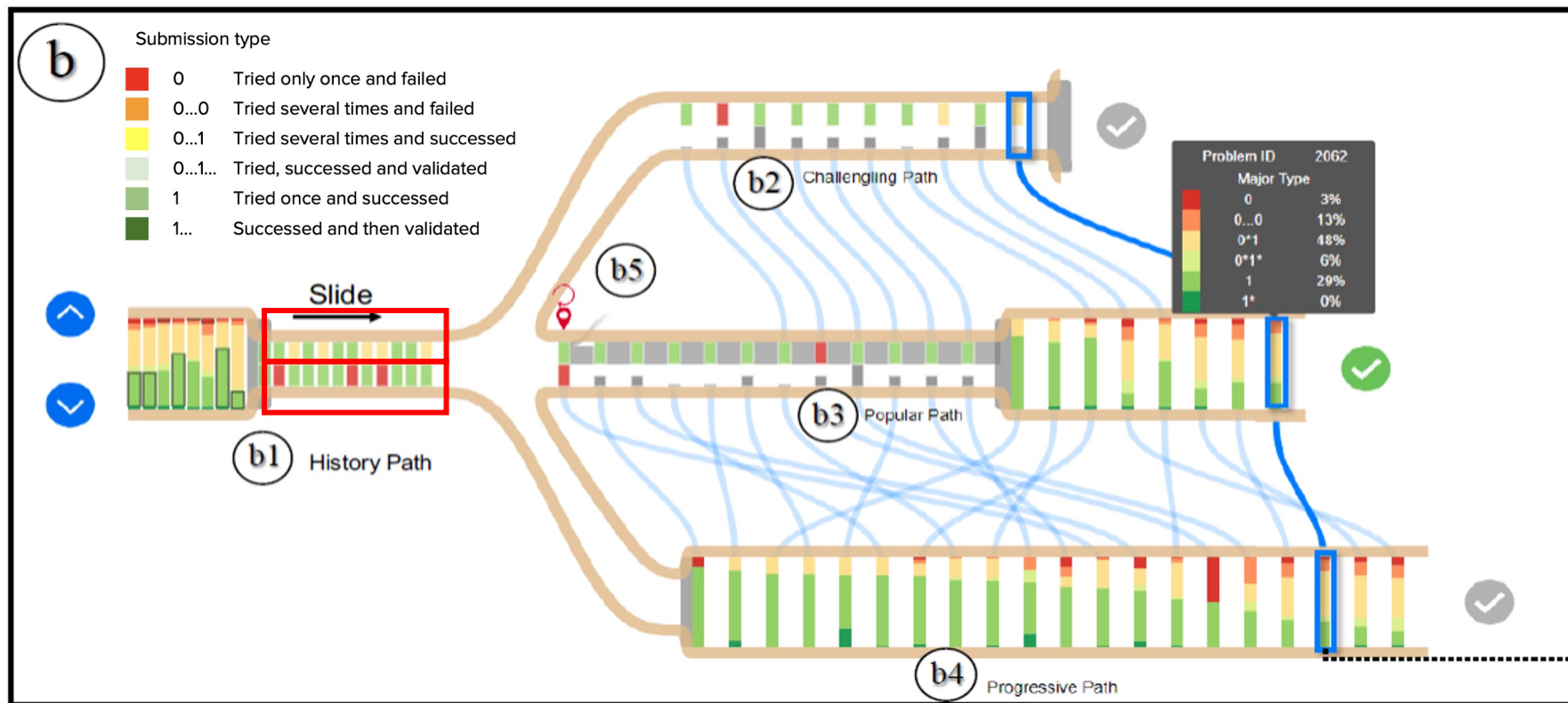
Visual Design: Learning Path View



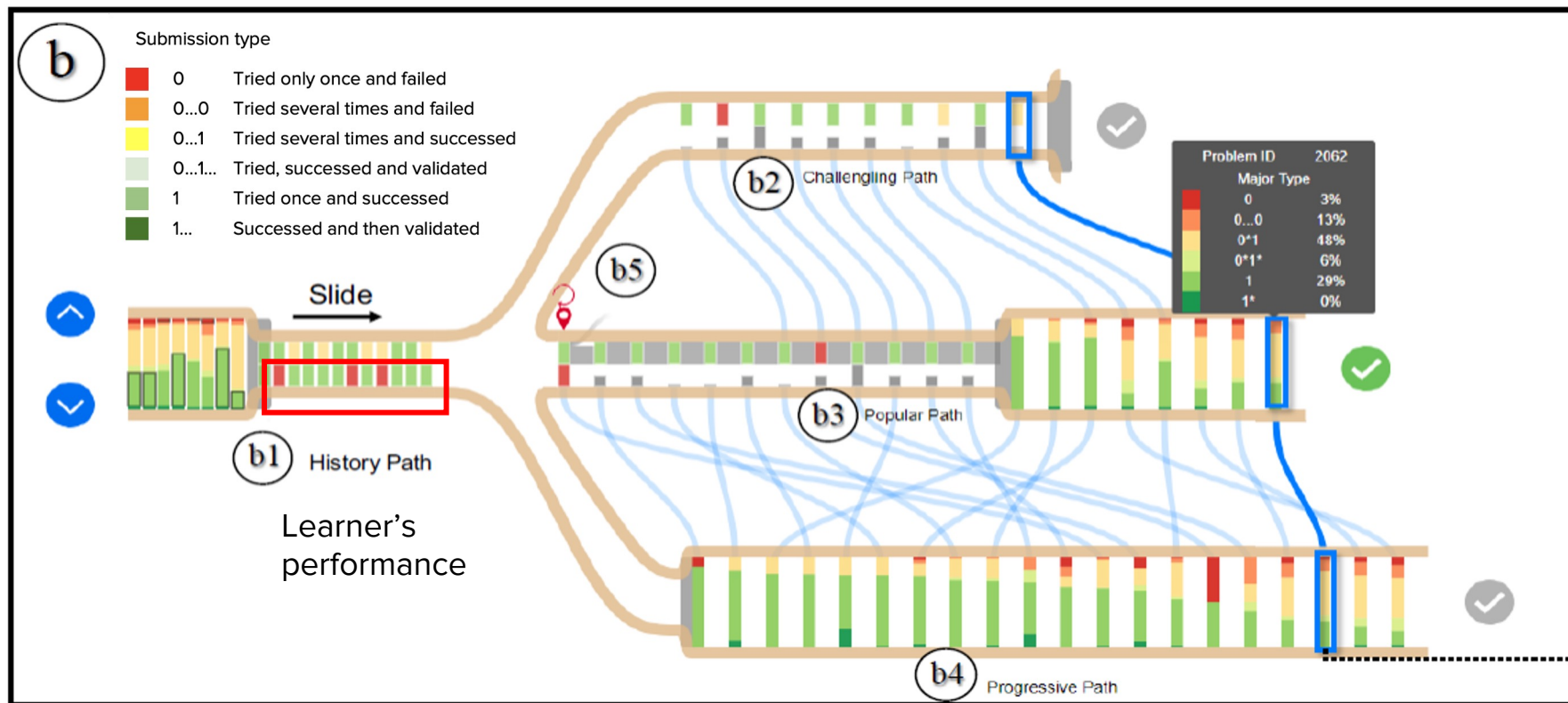
Visual Design: Learning Path View



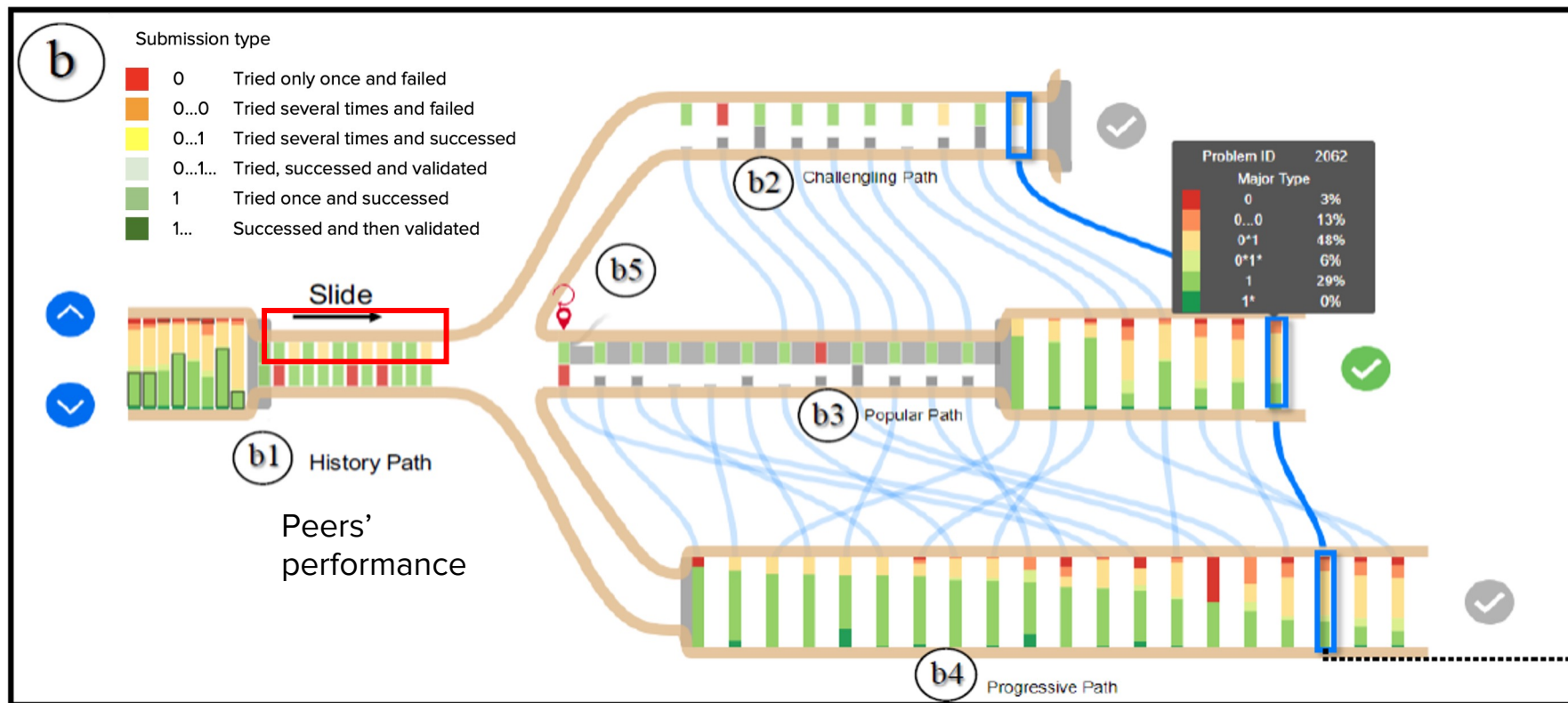
Visual Design: Learning Path View



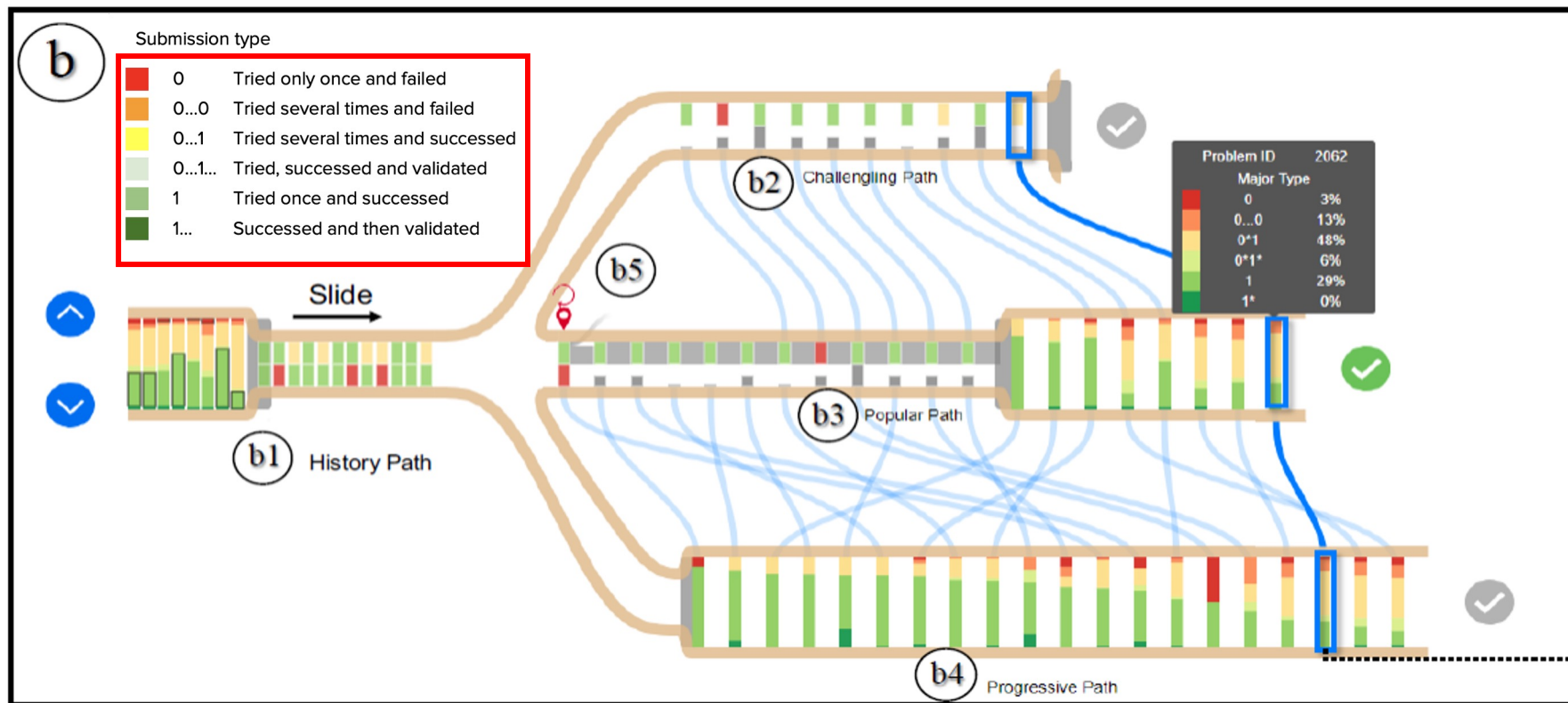
Visual Design: Learning Path View



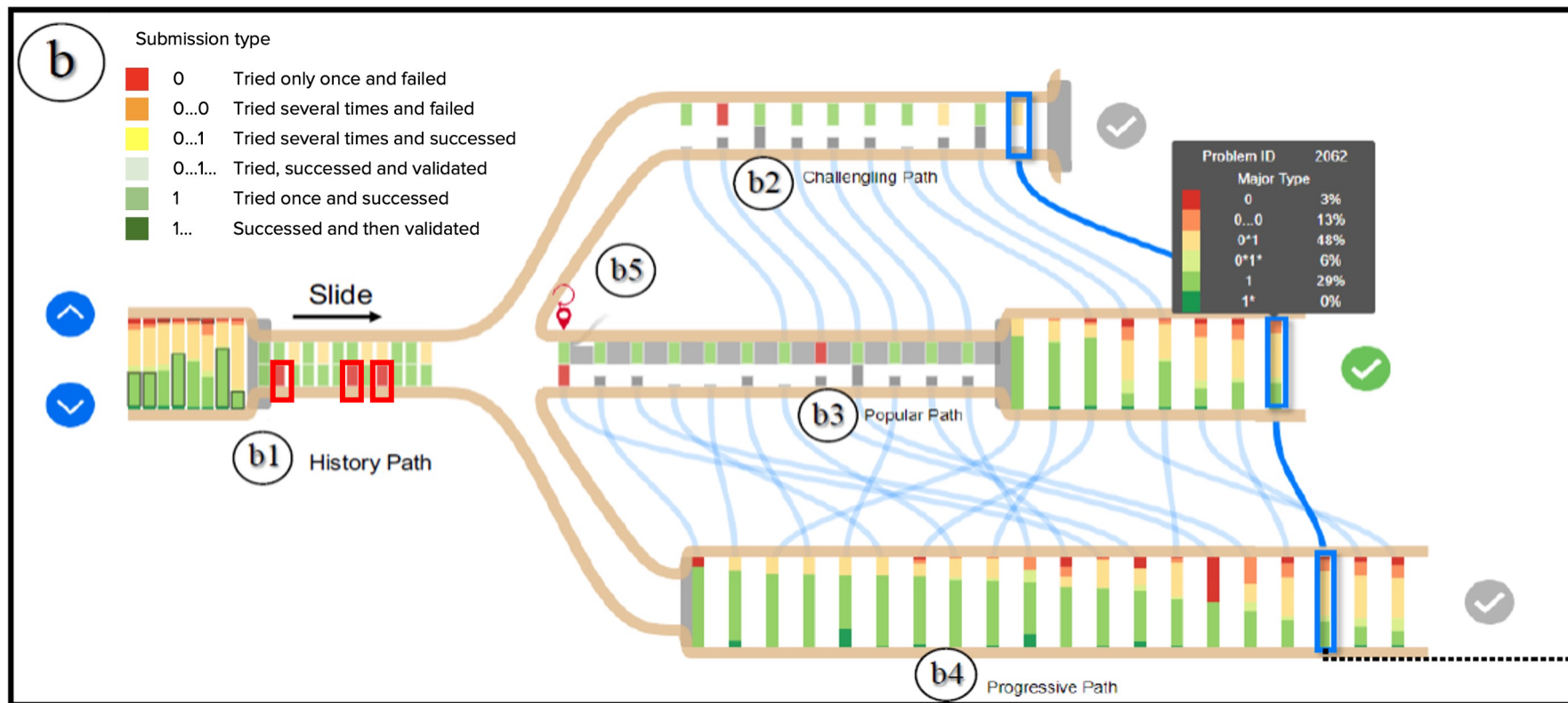
Visual Design: Learning Path View



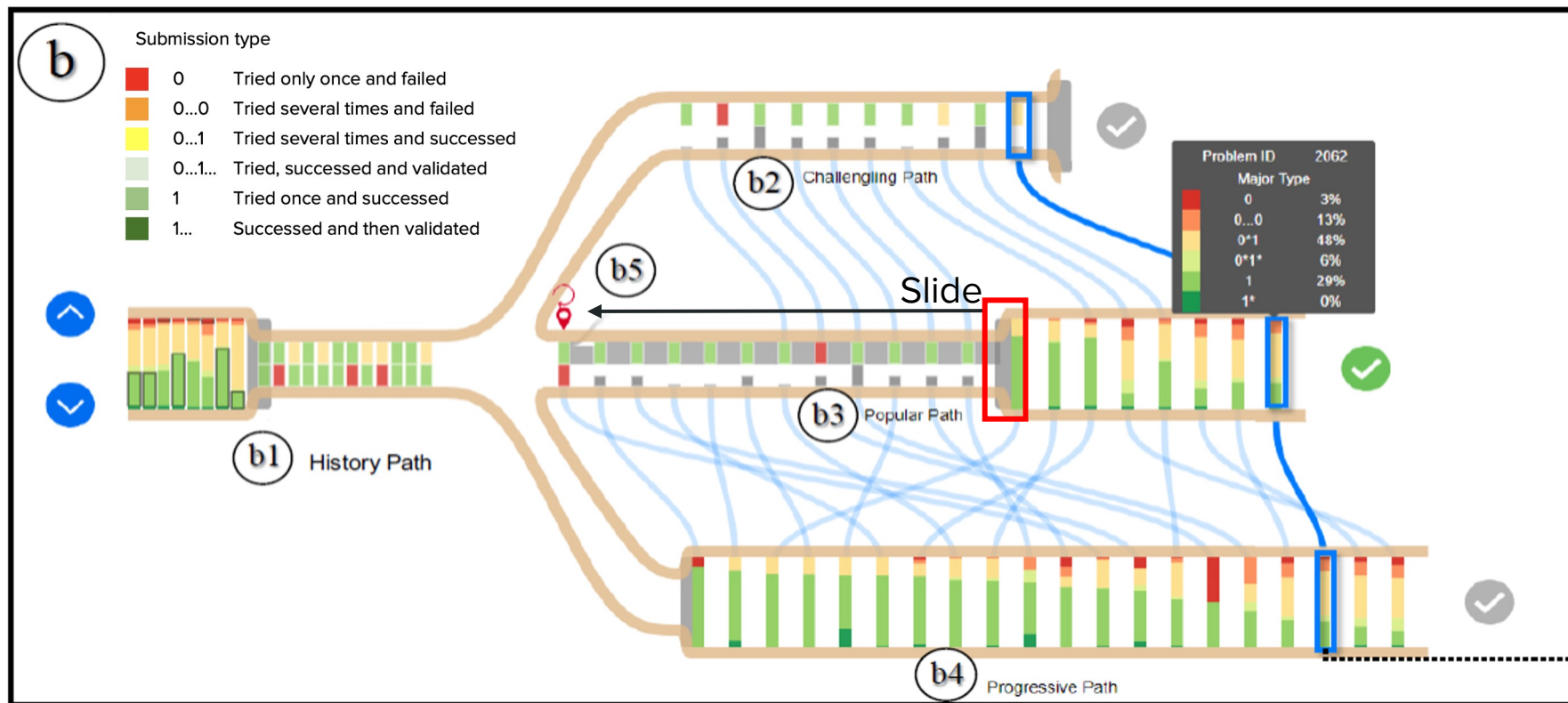
Visual Design: Learning Path View



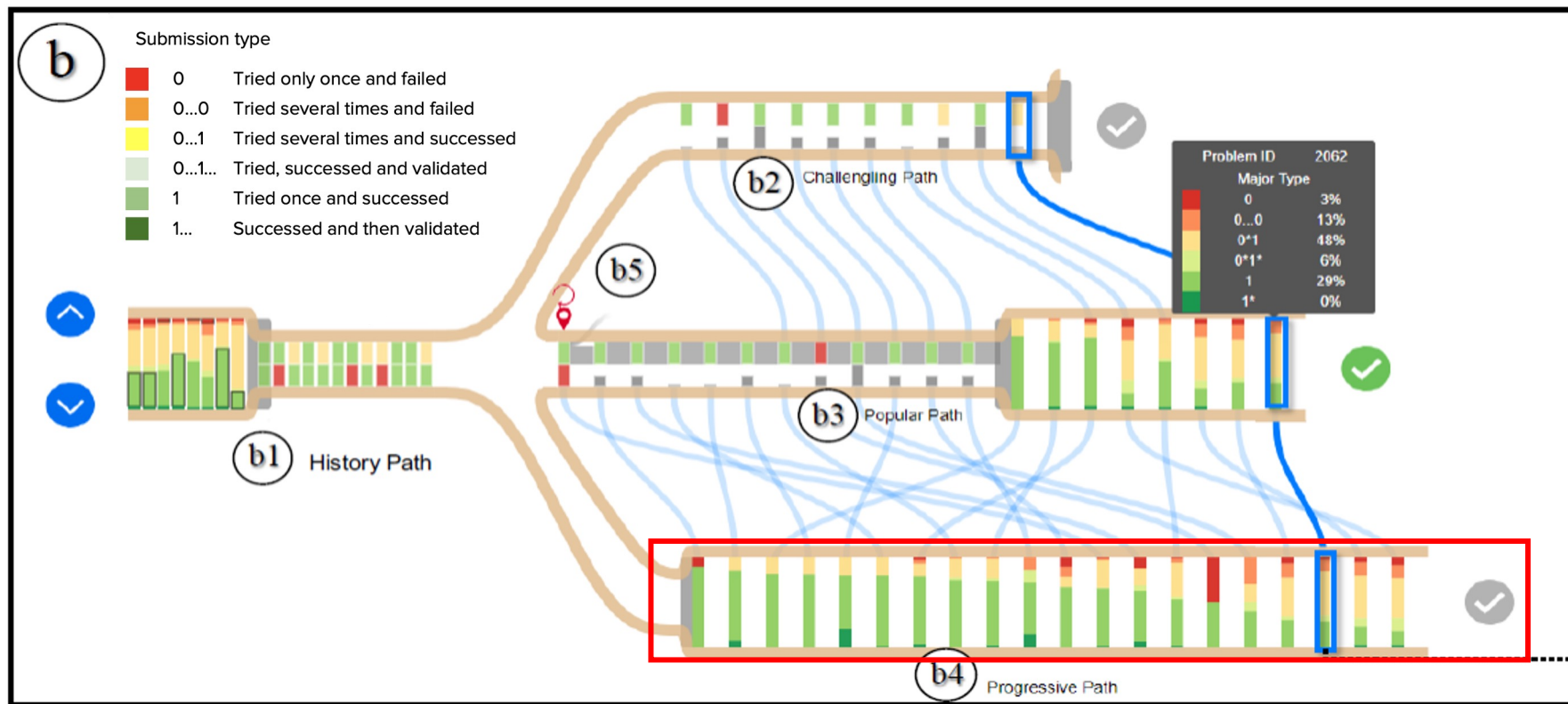
Visual Design: Learning Path View



Visual Design: Learning Path View



Visual Design: Learning Path View



Evaluation

RQ1: Is peer data useful?

RQ2: Is visualizing more suggestions helpful for planning compared to only one path?

RQ3: Does visualizing more suggestions using the proposed visualizations increase complexity?

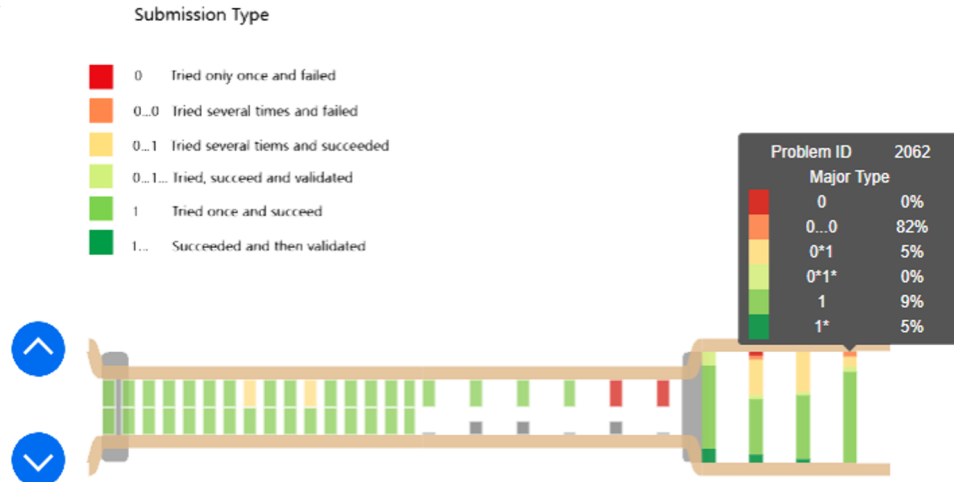
Evaluation: Controlled User Study

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15...33 34 35 36 37 38
39 40 41 42 43 44 45 46 47 48 49 50

Search: In

Pro. ID	Problem Title	Ratio(Accepted/Submissions)
1000	A + B Problem	30.56%(240770/787844)
1001	Sum Problem	25.38%(143110/563922)
1002	A + B Problem II	19.47%(84152/432201)
1003	Max Sum	23.76%(70413/296345)
1004	Let the Balloon Rise	39.72%(59043/148661)
1005	Number Sequence	25.25%(51499/203970)
1006	Tick and Tick	26.73%(6080/22750)
1007	Quoit Design	26.52%(17197/64856)
1008	Elevator	54.79%(46878/85565)
1009	FatMouse' Trade	34.85%(33070/94883)
1010	Tempter of the Bone	26.68%(39786/149139)

Baseline system (List View)



Primitive PeerLens (Only provide one path)

18 CS students :

- determine the starting question under a specific learning scenario
- 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.

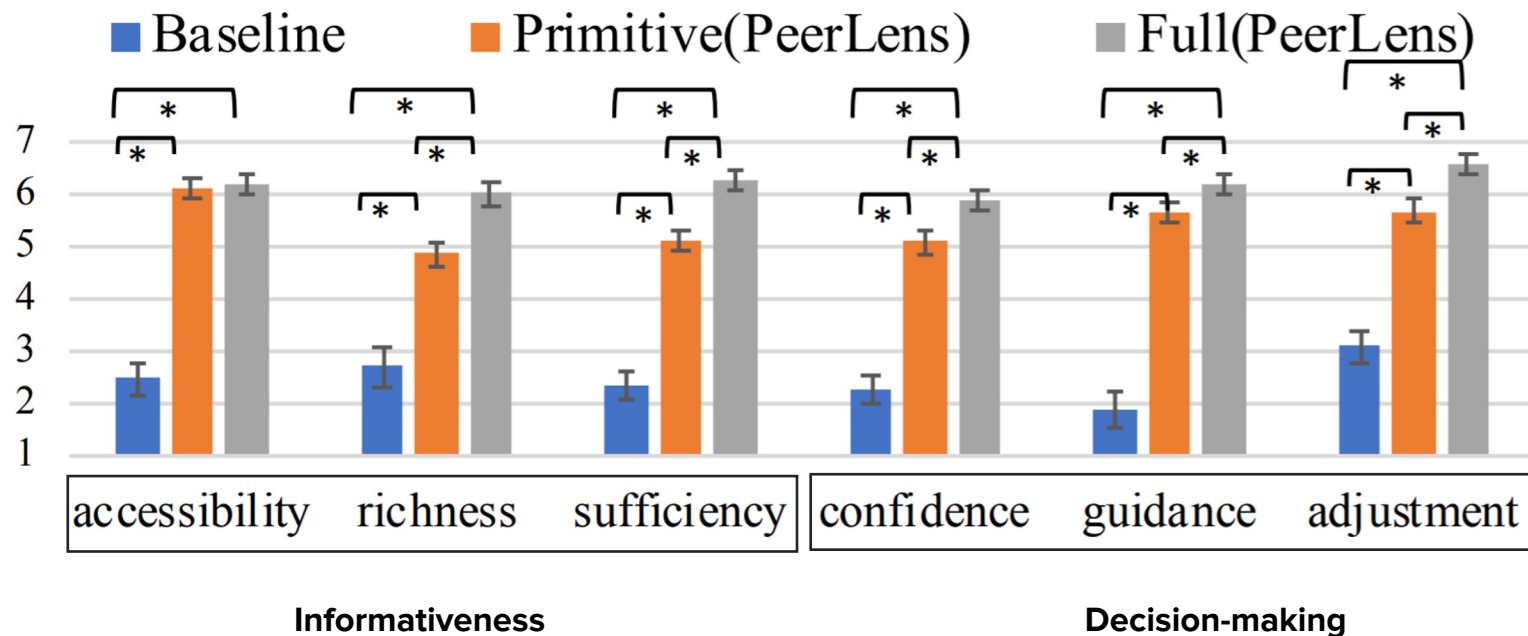
Visual design

Q7	The learning path design is intuitive.
Q8	The learning path design helps me understand the suggested path.

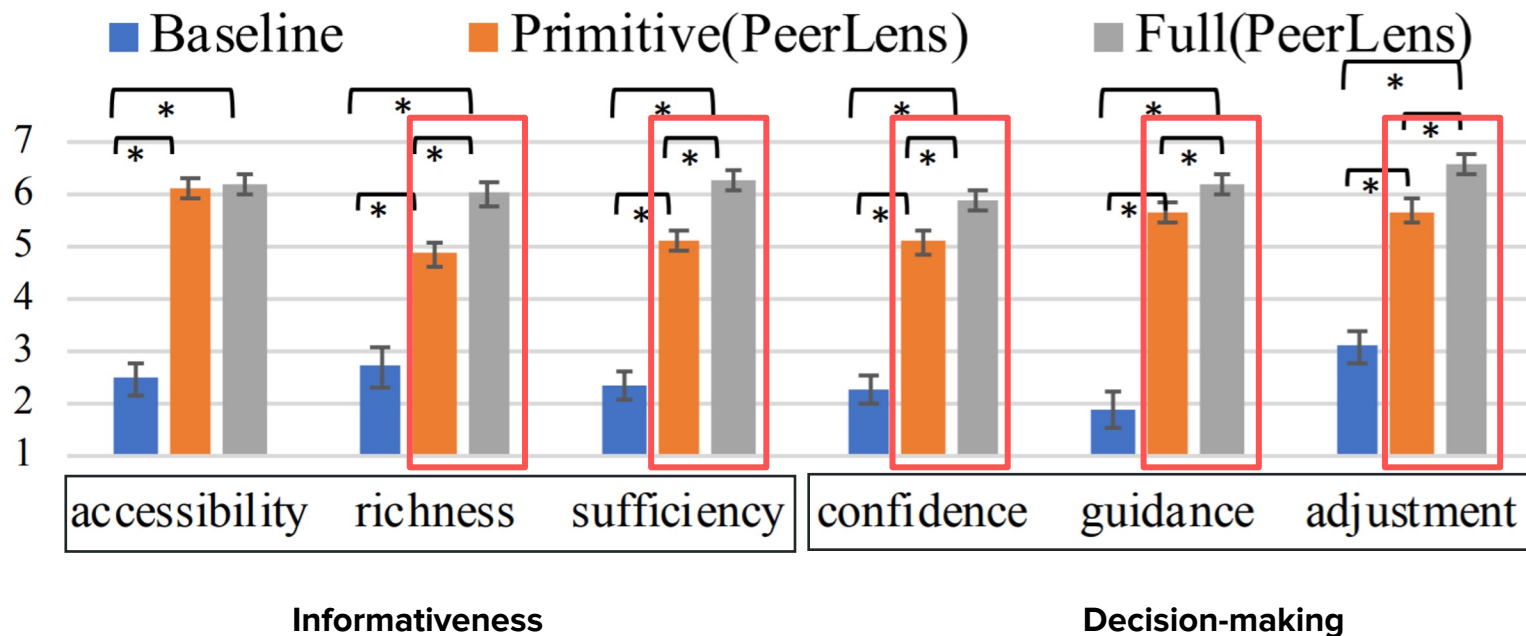
System Usability

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.

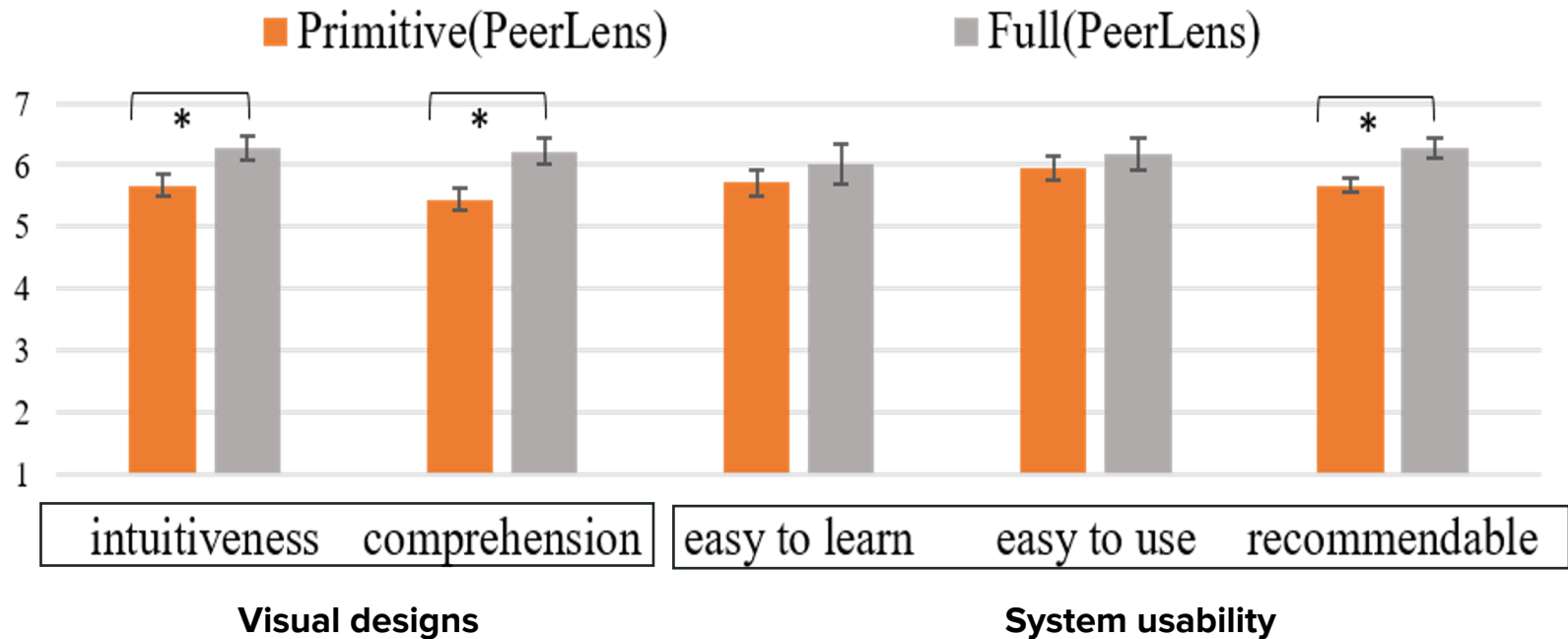
Results: Using peer data is useful (RQ1)



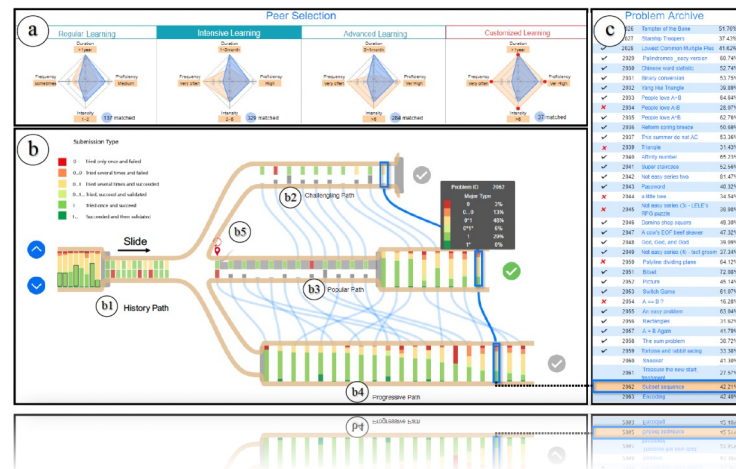
Results: Visualizing more suggestions is useful (RQ2)



Results: Visualizing more suggestions using the proposed visualizations did not increase the complexity (RQ3)



With **PeerLens**, we enable students to use **peer data** to plan their own learning path by **suggesting** and **visualizing multiple paths**.



Sequential event-based learning, e.g., other question pools, MOOCs, etc.



Educators (design loop)



Learners (learning loop)

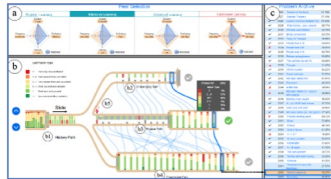
Planning

Performing
Data

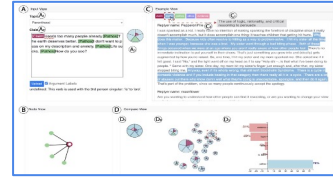
Reflecting

Understanding
and Analyzing

Redesigning



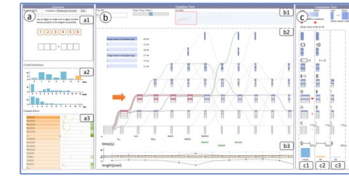
Peerlens (CHI 2019)



Persua (CSCW 2022)



RLens (L@S 2022)



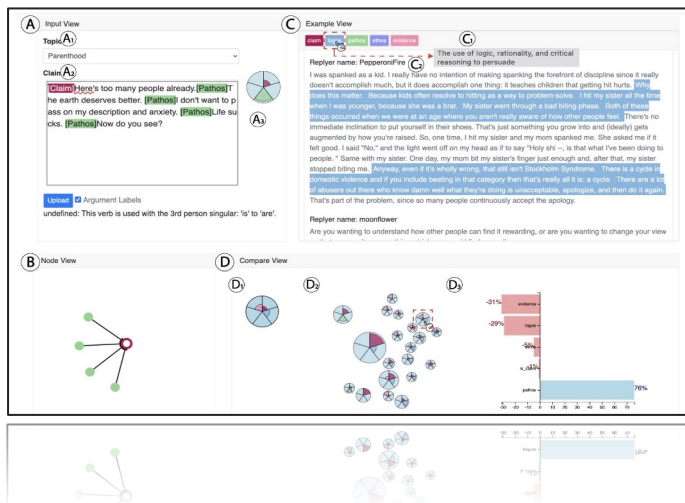
QLens (TVCG 2021)

How data can be used for learners to plan, **perform**, and reflect on their learning?

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

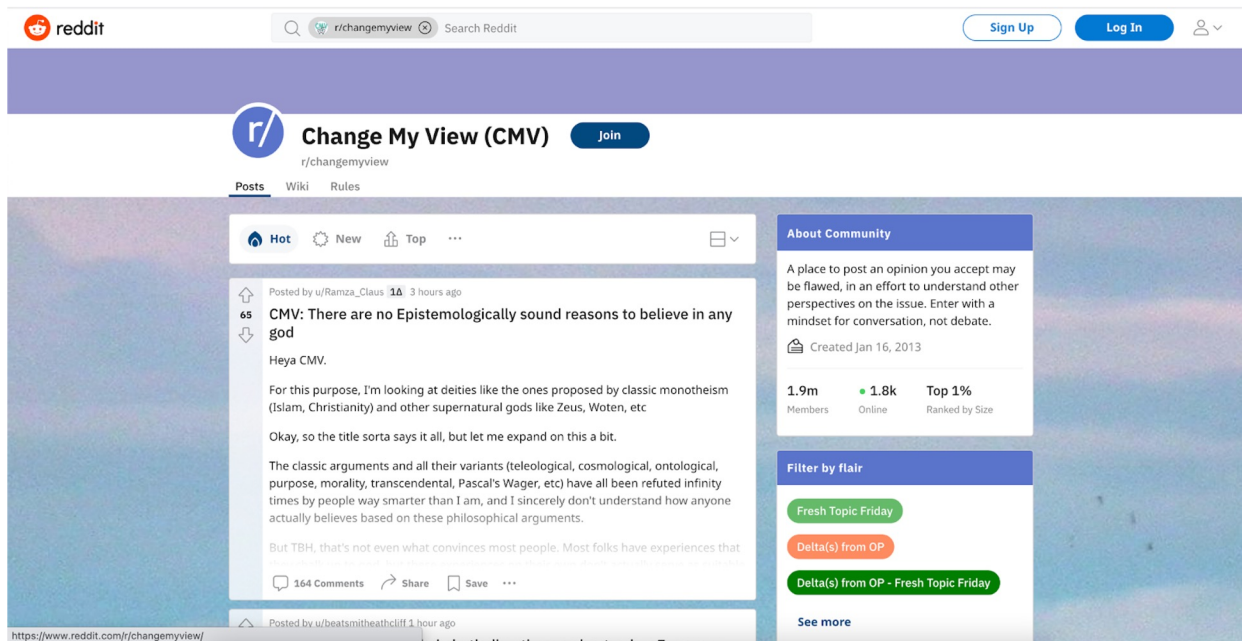
CSCW 2022



Background

- **Real-time adaptive feedback**
- **Critical thinking skills, e.g., arguments writing**, a major element of learning framework 2030 by the Organization for Economic Co-operation and Development (OECD)

Background



1.9m
Members

1.8k
Online

Top 1%
Ranked by Size

What are the difficulties for users to write persuasive arguments and what we can learn from the existing data?

Needs-finding Stage



A needs-finding survey with
123 online forum users

R1: Provide examples and support filtering by persuasive strategies

R2: Display and **compare** the **composition** of different persuasive strategies

R3: Show the logical structure of the arguments

R4: Offer visual augmented feedback for the arguments writing

Related Work

AL

Home

Write

Progress

Text editor

While writing, you can click the **Analyze** button to get feedback on your argumentation.

Click on **claims** to display them in the detailed view.

Die Idee ist im Grundsatz klar: Eine App, welche alles im Haushalt automatisiert, was möglich ist. Es wird wahrscheinlich immer mehr Services auch im Bezug auf den Haushalt geben. Entsprechend ist der Zeitpunkt und die Stossrichtig gut gewählt. Es gibt verschiedenste Haushaltsaufgaben, welche grossen Informationsaufwand oder Koordinationsaufwand mit sich bringen insbesondere wenn man diese durch ServiceAnbieter erledigen lassen möchte. Es ist mir nicht klar, wie dem Kunden gute Dienstleistungen garantiert werden können. Dies ist vor allem wahr, wenn wir von einer multinationalen oder gar weltweit nutzbaren Plattform sprechen. Menschen haben häufig Vorlieben für Dinge, die sie bereits kennen. So weiss ich aus Erfahrung, dass man sich bei Handwerkern gerne jene aussucht, von denen man schon Gutes gehört oder mit welchen man im Idealfall schon gute Erfahrungen gemacht hat. Es ist mir nicht klar, wie dies über die genannte App passieren würde vor allem wenn das geografische Einzugsgebiet gross ist. Zudem wäre aus meiner Sicht Transparenz bezüglich des genauen Providers sehr wichtig. Dies wäre mein Vorschlag für dich.

Words: 167

Analyze

Explanation

Your personal argumentation learning dashboard

Our algorithms found the following feedback (to know how, click [here](#))

An overview of your argumentation structure (claims & premises):

Click on **claims** to display them in the detailed view.

Argumentative [89%]
Non-Argumentative [11%]

Readability 64% [Details](#)

Coherence 72% [Details](#)

Persuasiveness 75% [Details](#)

Detailed feedback on the structure of your claims (What are claims and premises?):

Novice Advanced Competent Proficient Expert

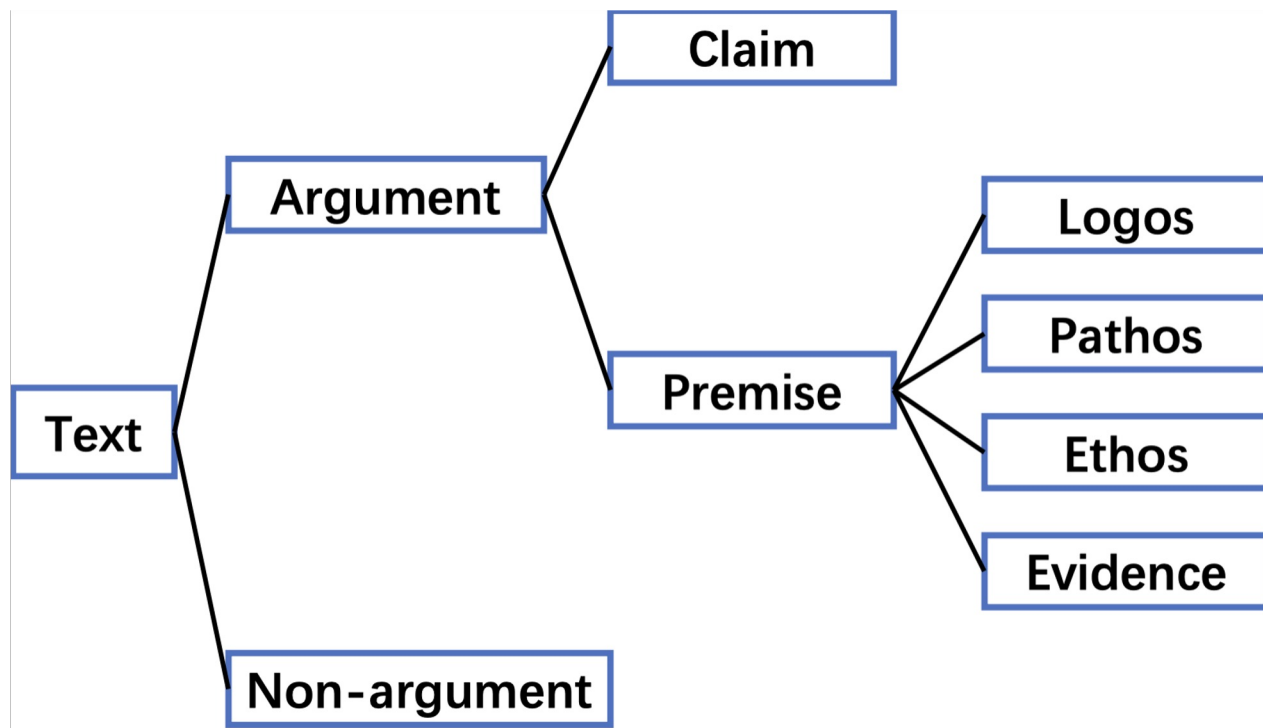
Your argumentation improved by 24 % compared to the last exercise.

Lacking guidance on persuasive strategies

Taxonomy of Argument Components



Semi-structured interviews with **five domain experts** on debating competitions

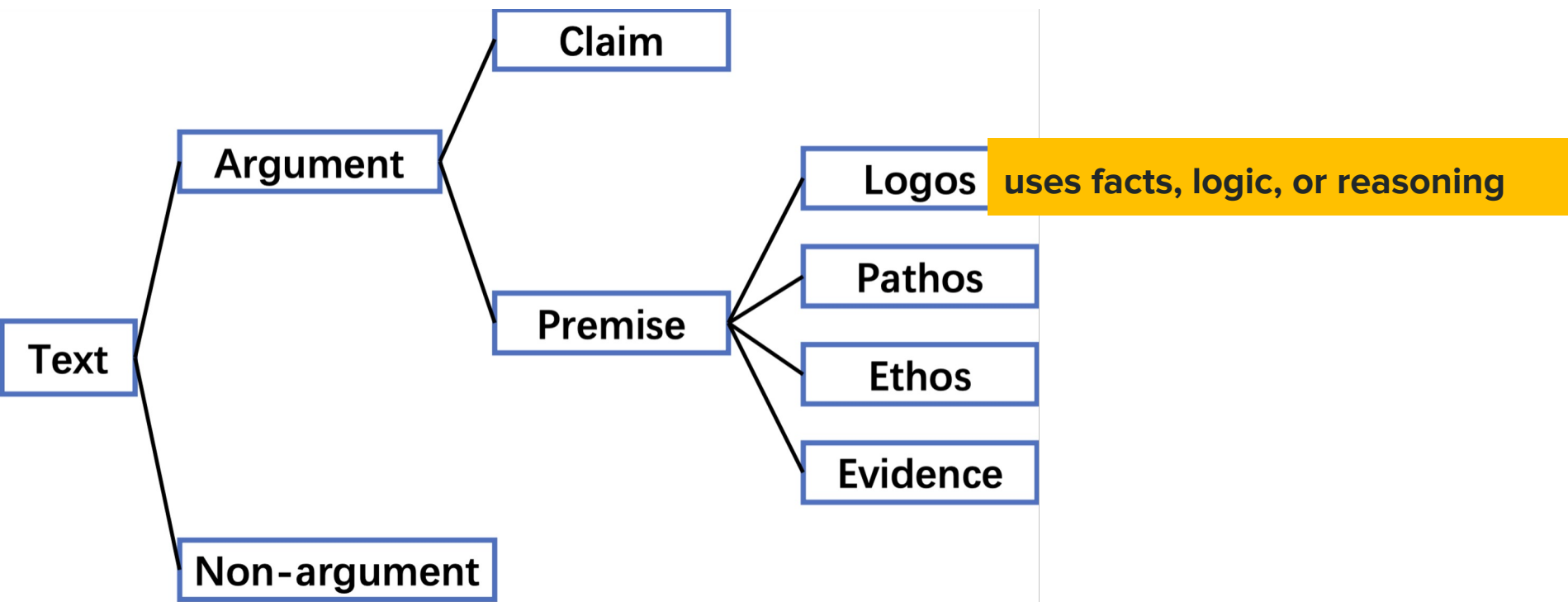


Classical persuasive strategies of Aristotle, and Rapp, 2002 and Carlile et al., 2018

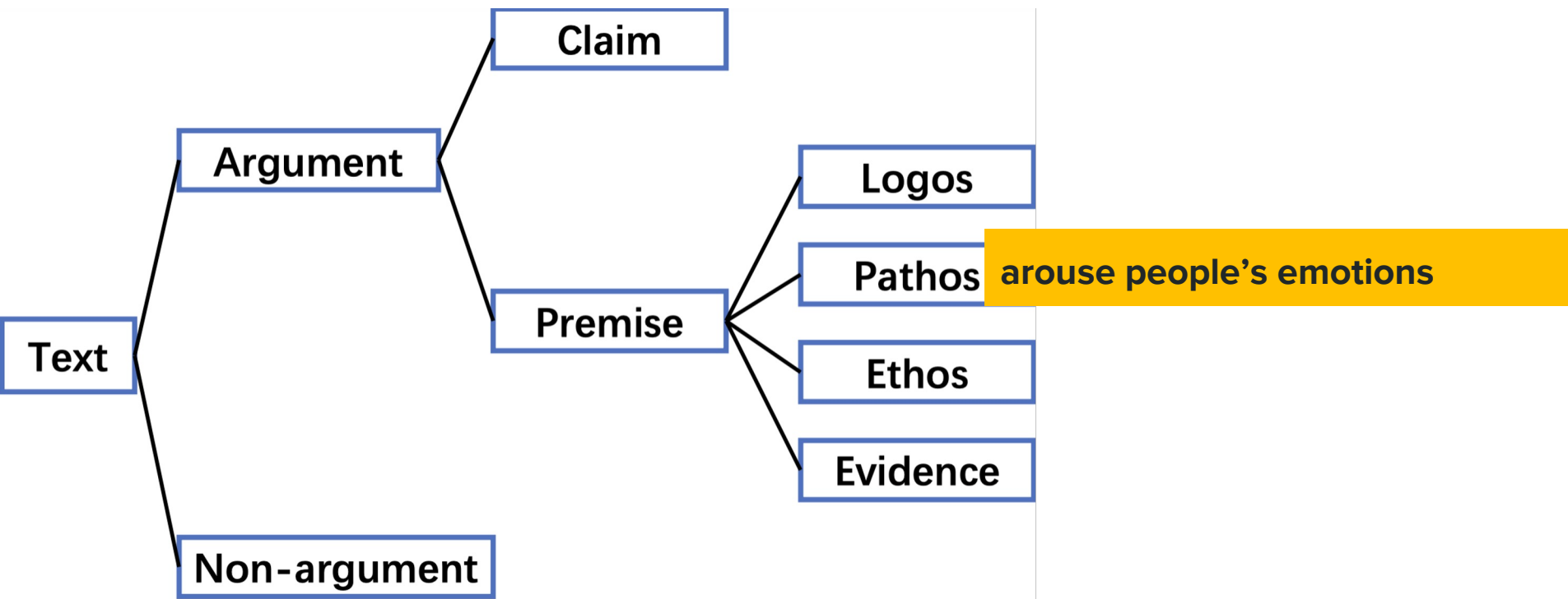
Rapp, Christof. "Aristotle's rhetoric." (2002).

Carlile, Winston, et al. "Give me more feedback: Annotating argument persuasiveness and related attributes in student essays." *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2018.

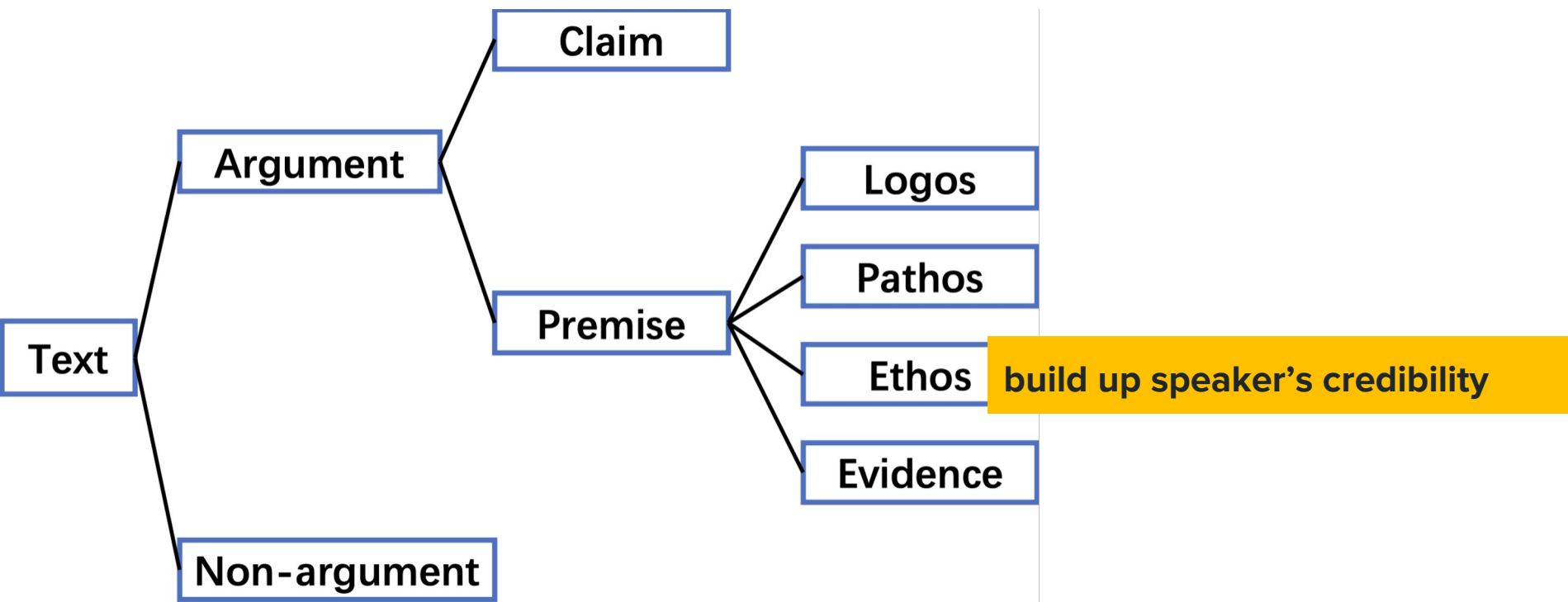
Taxonomy of Argument Components



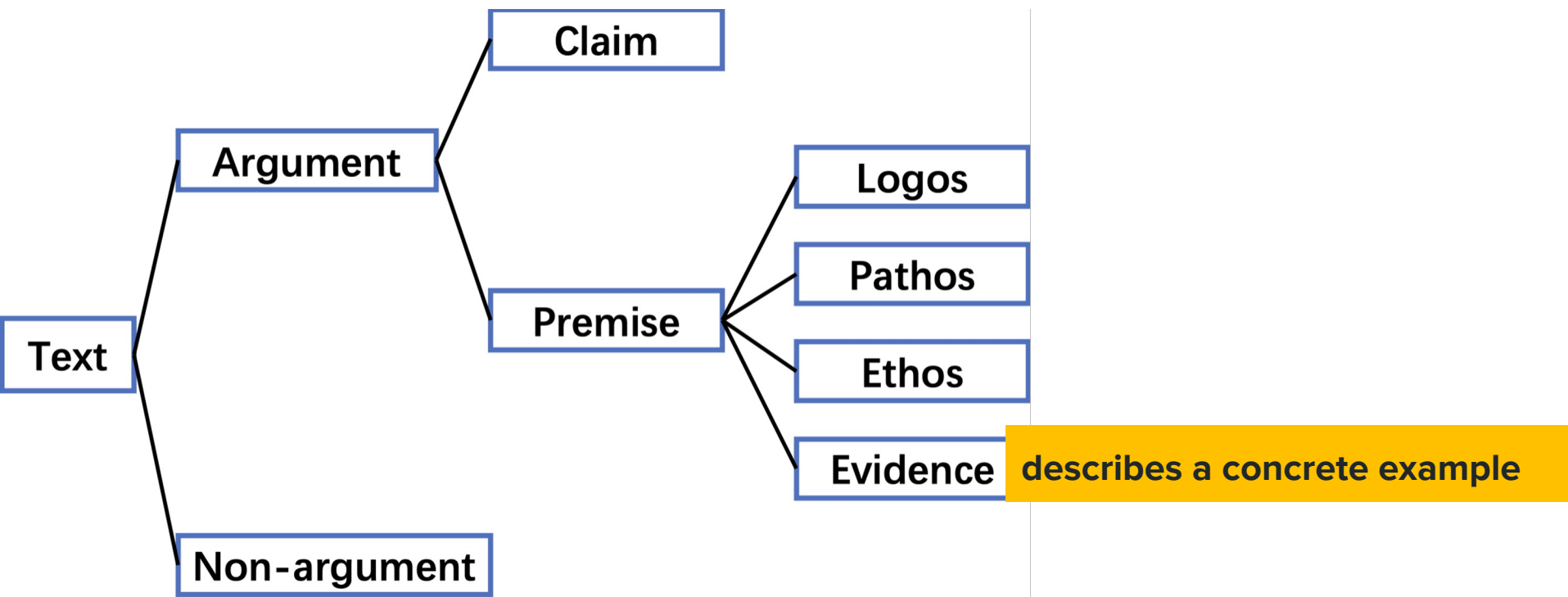
Taxonomy of Argument Components



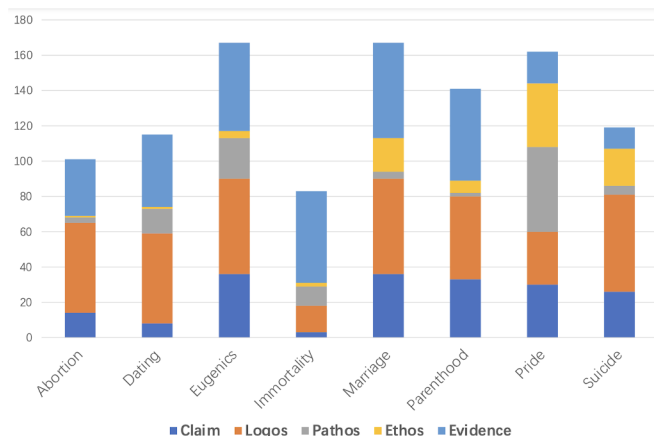
Taxonomy of Argument Components



Taxonomy of Argument Components



Persuasive Strategies Mining



164 discussion threads with 1269 sentences coving **eight** topics and

1. **Converting the sentences into high-dimensional vector representations:** BERT language model
2. **Is claim or premise or non-argument?** a multi-class classification task
3. **Does a premise support a claim?** a binary classification task
4. **What are the persuasive strategies applied?** a multi-label classification

Persua

A Input View

Topic **A₁**

Parenthood

Claim **A₂**

[Claim] Here's too many people already. [Pathos] The earth deserves better. [Pathos] I don't want to pass on my depression and anxiety. [Pathos] Life sucks. [Pathos] Now do you see?

Upload ☒ Argument Labels

undefined: This verb is used with the 3rd person singular: 'is' to 'are'.

A₃

C Example View

claim logos pathos ethos evidence

Replier name: **PepperoniFire**

The use of logic, rationality, and critical reasoning to persuade

I was spanked as a kid. I really have no intention of making spanking the forefront of discipline since it really doesn't accomplish much, but it does accomplish one thing: it teaches children that getting hit hurts. Why does this matter. Because kids often resolve to hitting as a way to problem-solve. I hit my sister all the time when I was younger, because she was a brat. My sister went through a bad biting phase. Both of these things occurred when we were at an age where you aren't really aware of how other people feel. There's no immediate inclination to put yourself in their shoes. That's just something you grow into and (ideally) gets augmented by how you're raised. So, one time, I hit my sister and my mom spanked me. She asked me if it felt good. I said "No," and the light went off on my head as if to say "Holy shi --, is that what I've been doing to people." Same with my sister. One day, my mom bit my sister's finger just enough and, after that, my sister stopped biting me. Anyway, even if it's wholly wrong, that still isn't Stockholm Syndrome. There is a cycle in domestic violence and if you include beating in that category then that's really all it is: a cycle. There are a lot of abusers out there who know damn well what they're doing is unacceptable, apologize, and then do it again. That's part of the problem, since so many people continuously accept the apology.

Replier name: moonflower

Are you wanting to understand how other people can find it rewarding, or are you wanting to change your view

B Node View

D Compare View

D₁

D₂

D₃

Label	D ₁	D ₂	D ₃
evidence	-31%	-29%	-5%
logos	-29%	-5%	-1%
ethos	-5%	-1%	1%
is_claim	-1%	1%	76%
pathos	76%	76%	76%

Input View

A

Input View

C

A₁

Topic

Parenthood

▼

A₂

Claim

[Claim]

Here's too many people already.

[Pathos]

The earth deserves better.

[Pathos]

I don't want to pass on my depression and anxiety.


[Pathos]

Life sucks.

[Pathos]

Now do you see?

A₃



Upload

☒ Argument Labels

undefined: This verb is used with the 3rd person singular: 'is' to 'are'.

Input View

A

Input View

C

A₁

Topic

Parenthood

▼

A₂

Claim

[Claim]

Here's too many people already.

[Pathos]

The earth deserves better.

[Pathos]

I don't want to pass on my depression and anxiety.


[Pathos]

Life sucks.

[Pathos]

Now do you see?

A₃

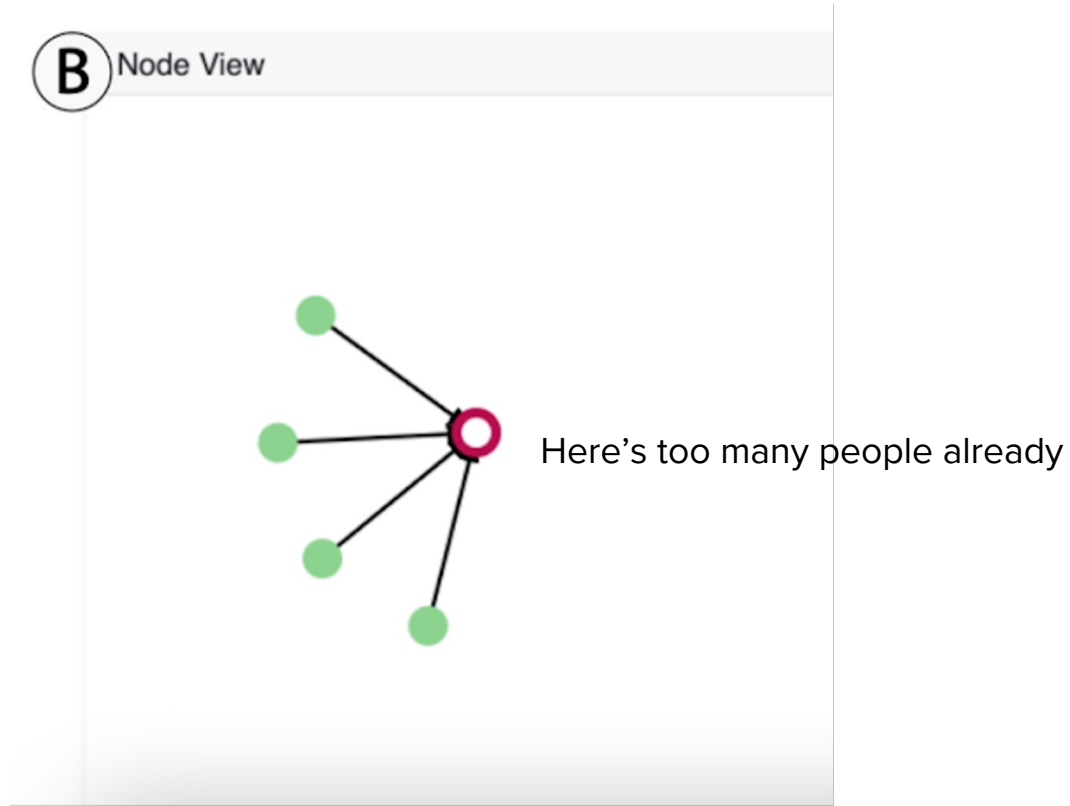


Upload

☒ Argument Labels

undefined: This verb is used with the 3rd person singular: 'is' to 'are'.

Node View



Example View

C Example View

claim logic pathos ethos evidence

G₁

The use of logic, rationality, and critical reasoning to persuade

Replier name: PepperoniFire **C₂**

I was spanked as a kid. I really have no intention of making spanking the forefront of discipline since it really doesn't accomplish much, but it does accomplish one thing: it teaches children that getting hit hurts. Why does this matter. Because kids often resolve to hitting as a way to problem-solve. I hit my sister all the time when I was younger, because she was a brat. My sister went through a bad biting phase. Both of these things occurred when we were at an age where you aren't really aware of how other people feel. There's no immediate inclination to put yourself in their shoes. That's just something you grow into and (ideally) gets augmented by how you're raised. So, one time, I hit my sister and my mom spanked me. She asked me if it felt good. I said "No," and the light went off on my head as if to say "Holy shi --, is that what I've been doing to people." Same with my sister. One day, my mom bit my sister's finger just enough and, after that, my sister stopped biting me. Anyway, even if it's wholly wrong, that still isn't Stockholm Syndrome. There is a cycle in domestic violence and if you include beating in that category then that's really all it is: a cycle. There are a lot of abusers out there who know damn well what they're doing is unacceptable, apologize, and then do it again. That's part of the problem, since so many people continuously accept the apology.

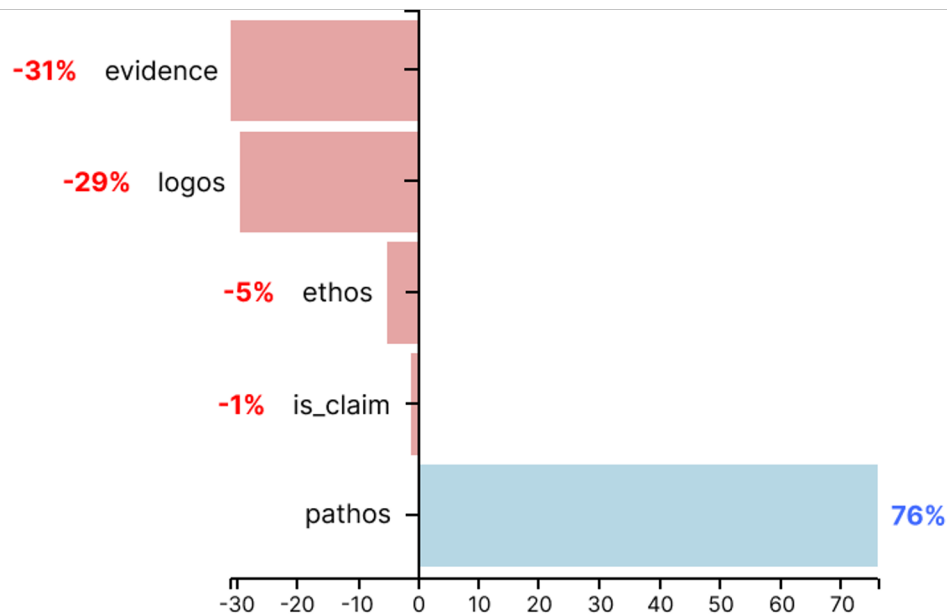
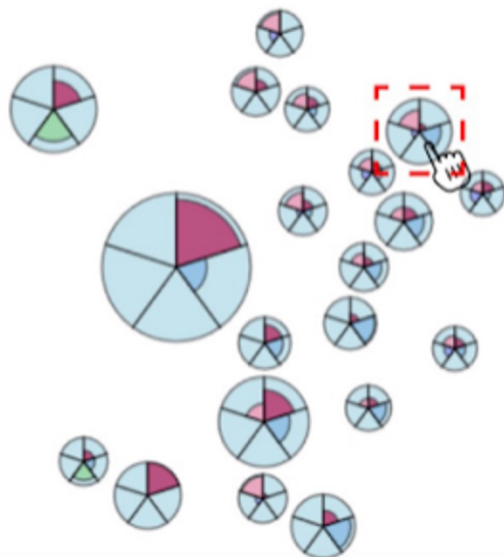
Replier name: moonflower

Are you wanting to understand how other people can find it rewarding, or are you wanting to change your view

Compare View

D₂

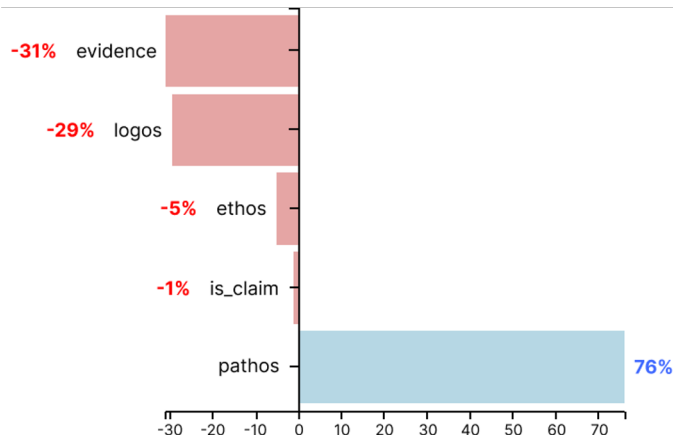
D₃



Case Study

Before:

[Claim] Here's too many people already. [Pathos] The earth deserves better. [Pathos] I don't want to pass on my depression and anxiety. [Pathos] Life sucks. [Pathos] Now do you see?



C Example View

claim logos pathos ethos evidence **C₁**

Replier name: PepperoniFire **C₂**

The use of logic, rationality, and critical reasoning to persuade

I was spanked as a kid. I really have no intention of making spanking the forefront of discipline since it really doesn't accomplish much, but it does accomplish one thing: it teaches children that getting hit hurts. Why does this matter. Because kids often resolve to hitting as a way to problem-solve. I hit my sister all the time when I was younger, because she was a brat. My sister went through a bad biting phase. Both of these things occurred when we were at an age where you aren't really aware of how other people feel. There's no immediate inclination to put yourself in their shoes. That's just something you grow into and (ideally) gets augmented by how you're raised. So, one time, I hit my sister and my mom spanked me. She asked me if it felt good. I said "No," and the light went off on my head as if to say "Holy shi --, is that what I've been doing to people." Same with my sister. One day, my mom bit my sister's finger just enough and, after that, my sister stopped biting me. Anyway, even if it's wholly wrong, that still isn't Stockholm Syndrome. There is a cycle in domestic violence and if you include beating in that category then that's really all it is: a cycle. There are a lot of abusers out there who know damn well what they're doing is unacceptable, apologize, and then do it again. That's part of the problem, since so many people continuously accept the apology.

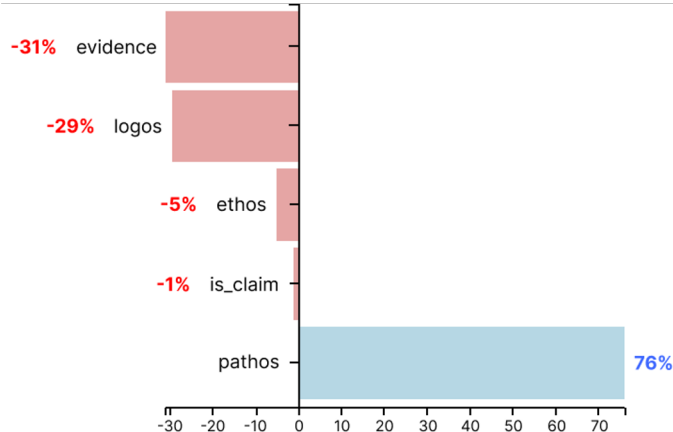
Replier name: moonflower

Are you wanting to understand how other people can find it rewarding, or are you wanting to change your view

Case Study

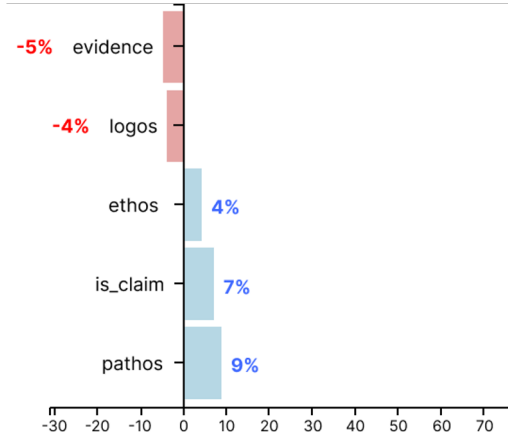
Before:

[Claim] Here's too many people already. [Pathos] The earth deserves better. [Pathos] I don't want to pass on my depression and anxiety. [Pathos] Life sucks. [Pathos] Now do you see?

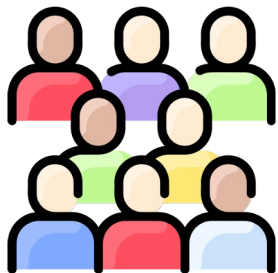


After:

[Claim] There is no need to have children. [Logos][Evidence] The world population is dramatically increasing, leading to more consumption of the natural resources. [Pathos][Evidence] In addition, I saw my uncle's children fight with each other. [Logos][Evidence] Some people even abuse children. [Claim] Thus, I disagree with having children.



User Study



36 participants (20 males, 16 females)
Baseline: 18 Persua: 18

Input Text

Claim:

Here's too many people already. The earth deserves better. I don't want to pass on my depression and anxiety. Life sucks. Now do you see?

Upload

ERROR undefined: This verb is used with the infinitive: 'to better', 'to well'.

Node View

Baseline (Wambsgans et al., CHI 2020)

A Input View

Topic: Parenthood

Claim: There's too many people already. The earth deserves better. I don't want to pass on my depression and anxiety. Life sucks. Now do you see?

Upload

Argument Labels

undefined: This verb is used with the 3rd person singular: 'is' to 'are'.

C Example View

Replyer name: PeppercornPm

The use of logic, rationality, and critical reasoning to persuade

I was spanked as a kid. I really have no intention of making spanking the forefront of discipline since it really doesn't accomplish much, but it does accomplish one thing: it teaches children that getting hit hurts. [They] does this matter. Because kids often resolve to hitting as a way to problem-solve. I hit my sister all the time when I was younger, because she was a brat. My sister went through a bad biting phase. Both of those things occurred during the worst of all for anyone who wants happy events of their child's period. There's no immediate indication to put yourself in their shoes. There's just something you grow into and (obviously) gets augmented by how you're raised. So, one time, I hit my sister and my mom spanked me. She asked me if it felt good. I said "No," and the light went off on my head as if to say "Holy shit -- is that what I've been doing to people." Same with my sister. One day, my mom bit my sister's finger just enough and, after that, my sister stopped biting me. Anyway, even if it's wholly wrong, that still isn't Stockholm Syndrome. There is a cycle in [different] violence and if you include beating in that category then that's really all it is: a cycle. There are a lot of abusers out there who know damn well what they're doing is unacceptable, apologize, and then do it again. That's part of the problem, since so many people continuously accept the apology.

Replyer name: moonflower

Are you wanting to understand how other people can find it rewarding, or are you wanting to change your view

B Node View

D Compare View

D1 D2 D3 D4

Persua

Tasks & Results: Submit more times and more persuasive

Write a paragraph about the topic “Abortion”.

Task 1 Try to persuade the person who published the following claim by taking the opposite side.
[Claim: I’m pro-life, and I believe that abortion is essentially murder.]

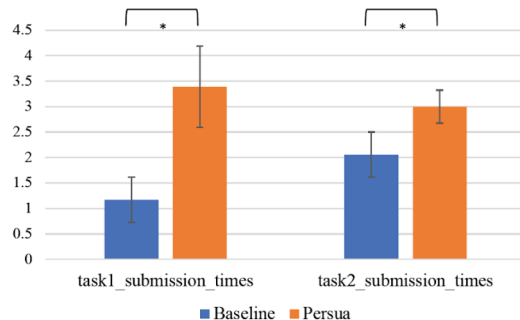
Refine a paragraph about the topic “Parenthood”:

“Here’s too many people already. The earth deserves better.

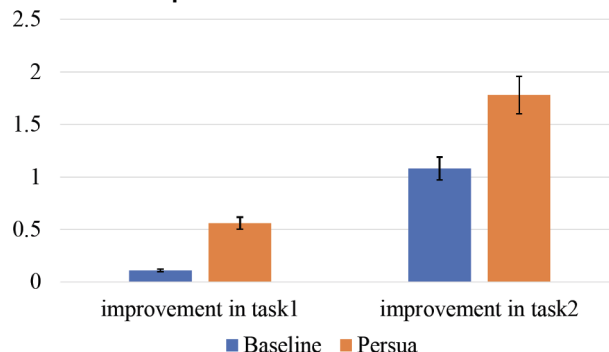
Task 2 *I don’t want to pass on my depression and anxiety. Life sucks. Now do you see?”*

Try to make it more convincing to persuade people who published the following claim.
[Claim: I don’t understand why people don’t want to have children.]

Submissions times on tasks

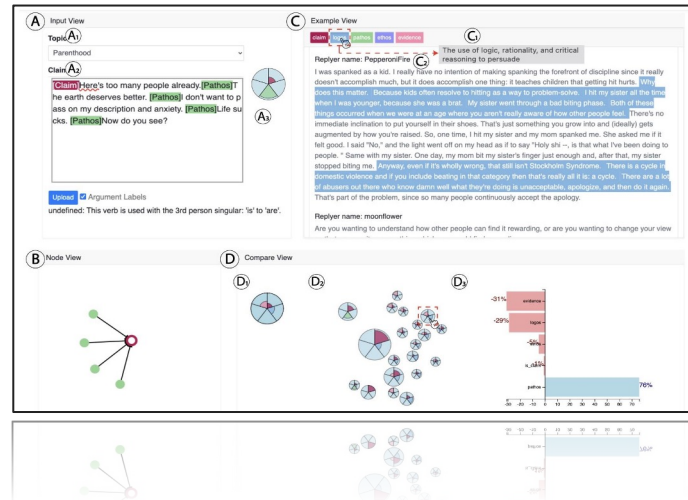


Improvements on tasks



Perusa detects and visualizes persuasive strategies differences from peer data to nudge learners and provide examples for them to refer to.

Component-based learning, e.g., analyze and visualize code structures and provide code examples to student



	Question	Result
Q1	How frequently do you usually use online discussion forums, e.g., Reddit, Baidu forum, Quora ? (1: never, 5: very often)?	2 14 8 59 40
Q2	How often do you actively express opinions in the discussion on hot topics in online discussion forums? (1: never, 5: very often)	19 14 48 38 18
Q3	How often are you worried that the content of the published text is not convincing enough to convince others or get recognition? (1: never, 5: very often)	9 18 39 50 39
Q4	How often do you encounter situations where you need to express your views and persuade others by publishing text content? (1: never, 5: very often)	3 18 40 34 22
Q5	How much do you want to improve your text persuasiveness through learning so as to get more support/convince others? (1: not at all, 5: very much)	2 10 36 51 32

F1: Participants frequently use online discussion forums and tend to express their opinions online. It confirms that university students fit the behavioral profile of our target users. As shown in Table 1, nearly all participants used (121 out of 123) online discussion forums, such as Reddit³ and Quora⁴ (Q1: $Mean = 3.98, SD = 1.00$) and 89% people (110 out of 123) at least once expressed their opinions on hot topics in online discussion forums (Q2: $Mean = 3.33, SD = 1.11$).

F2: Participants want to improve the persuasiveness of their arguments online. According to the answers from participants of Q3-Q5, we found that many participants (118 out of 123) were at least once worried that their arguments are not convincing enough (Q3: $Mean = 3.73, SD = 1.05$), and 120 out of 123 participants at least once encountered the situations to persuade others (Q4: $Mean = 3.44, SD = 1.03$). Most of the participants (121 out of 123) wanted to improve text persuasiveness through learning to some degree (Q5: $Mean = 3.82, SD = 0.97$).

F3: Participants find it hard to come up with evidence to support their arguments. Many participants wrote that writing or finding appropriate examples to support their opinions and make arguments persuasive is most time-consuming and challenging. P46 (F, 20) explained the difficulty of finding proper examples: "I want to find some relevant statistics to persuade others, but it is hard

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](#) · [♡](#) 21 ·

[↪](#) 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 that helps users improve the persuasiveness of arguments in online discussions through a survey with 123 online forum users and interviews with five debating experts. To satisfy these design goals, we analyzed and built a labeled dataset of fine-grained persuasive strategies (i.e., logos, pathos, ethos, and evidence) in 164 arguments with high ratings on persuasiveness from



Educators (design loop)



Learners (learning loop)

Planning

Performing
Data

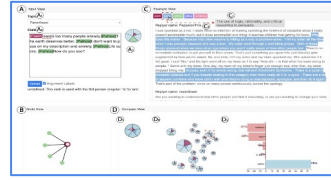
Reflecting

Understanding
and Analyzing

Redesigning



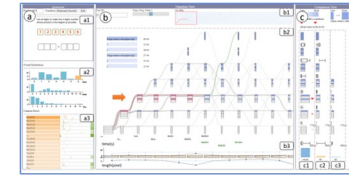
Peerlens (CHI 2019)



Persua (CSCW 2022)



RLens (L@S 2022)



QLens (TVCG 2021)

How data can be used for learners to plan, perform, and **reflect** on their learning?



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

Online language tutoring platforms (e.g., Cambly) are becoming increasingly popular.



CAMBLY



Soomgo



민병철유폰 3.0



SPICUS

당근영어

engoo

PAGODA 토크



Preply



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

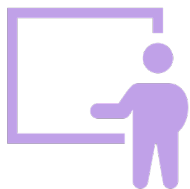
Distributed Tutorship

My previous work 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**.

(Xia et al., 2022)



Day 1/tutor 1



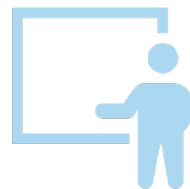
Day 2/tutor 2



Day 3/tutor3



Day 4/tutor 4



Day 5/tutor 5

Day ...



Distributed Tutorship

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

(Xia et al., 2022)



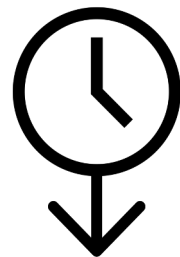
An online English tutoring platform. On Ringle, learners can choose tutors and class time for 1:1 online speaking sessions.



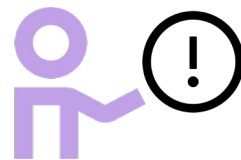
16 learners, who have learnt from more than one tutor.

Challenge: 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.



Grammar



Fluency



Vocabulary

NLP: Feedback uptake behaviors



Learners' corrective actions according to tutors' feedback

- Two apple -> Two apples
- “uh”
- “She always tries to **think positively.**” -> “She is always so **optimistic.**”

Visualization: Feedback uptake behaviors



RLens detects feedback uptake behaviors and visualizes the learning progress to keep consistency when learning under distributed tutorship.



Learning scenarios with subjective feedback from different teachers.



Our collaborator, Ringle, uses some of the algorithms and interface design on their platforms. The company now serves for **100,000+ users** over the world.



Educators (design loop)



Learners (learning loop)

Planning

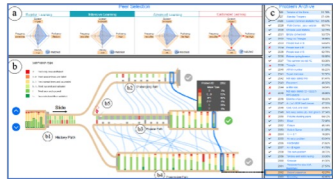
Performing

Data

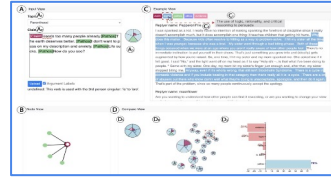
Reflecting

**Understanding
and Analyzing**

Redesigning



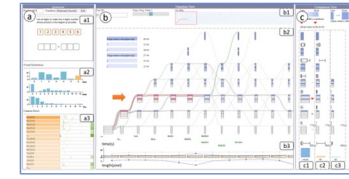
Peerlens (CHI 2019)



Persua (CSCW 2022)



RLens (L@S 2022)



QLens (TVCG 2021)

What if the existing learning materials and data online didn't cover different learners' needs?

What if the learners are too young to use the learning analytics?



Educators (design loop)



Learners (learning loop)

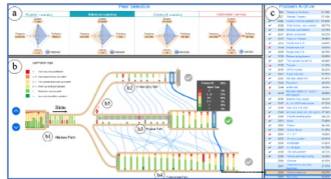
Planning

Performing
Data

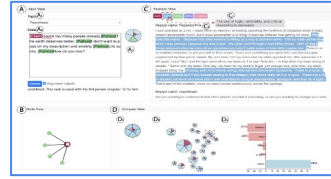
Reflecting

Understanding
and Analyzing

Redesigning



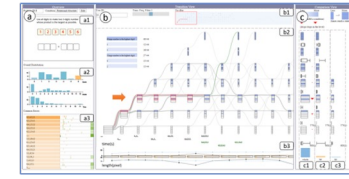
Peerlens (CHI 2019)



Persua (CSCW 2022)

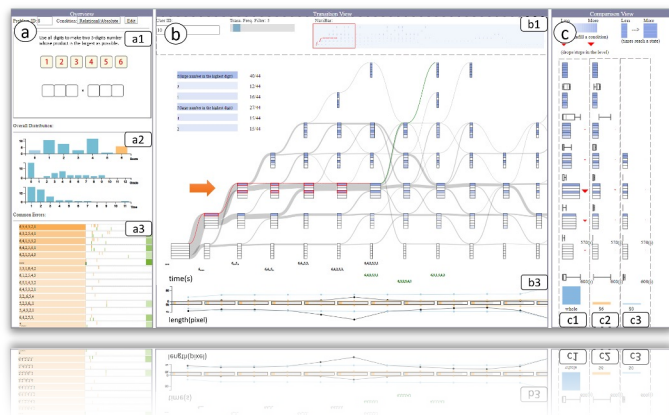


RLens (L@S 2022)



QLens (TVCG 2021)

How data can be used for educators to **improve the design of learning materials?**



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

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

TVCG 2021

100

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.



Motivation



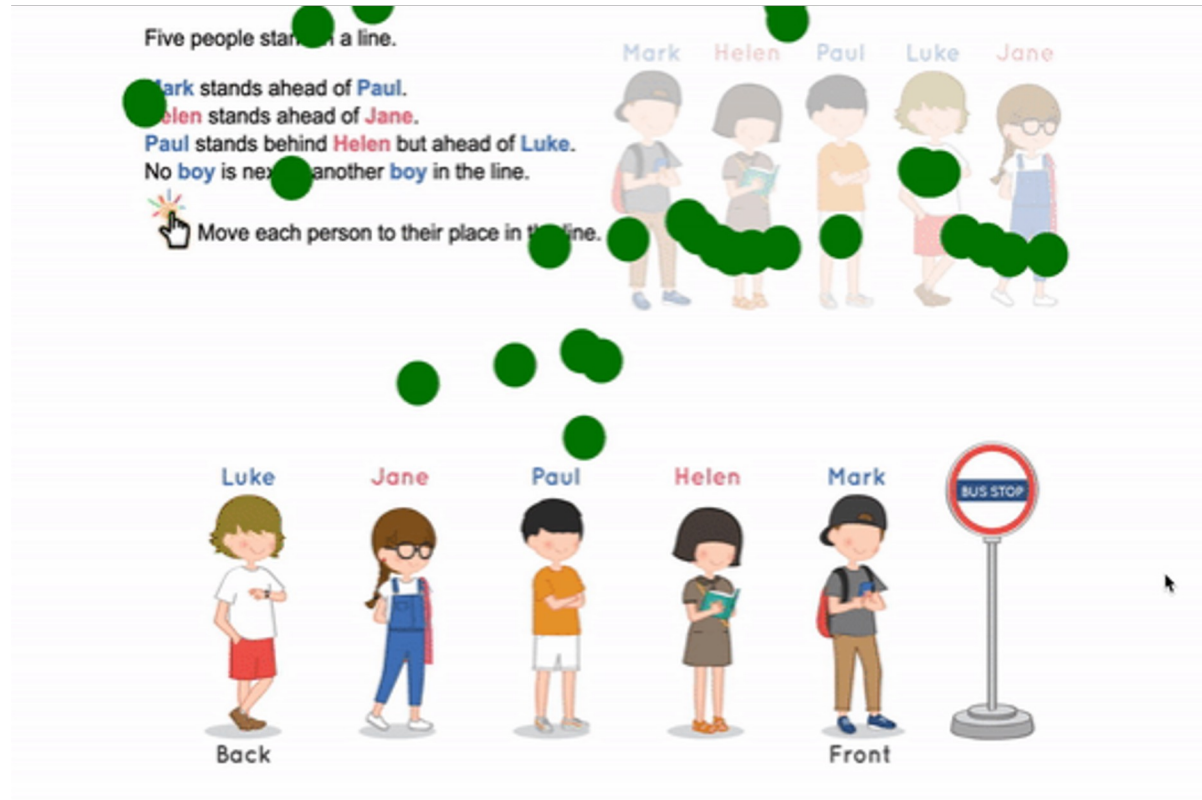
- Four domain experts

- Question designers (E1, E2)
- System developer (E3)
- Project manager (E4)

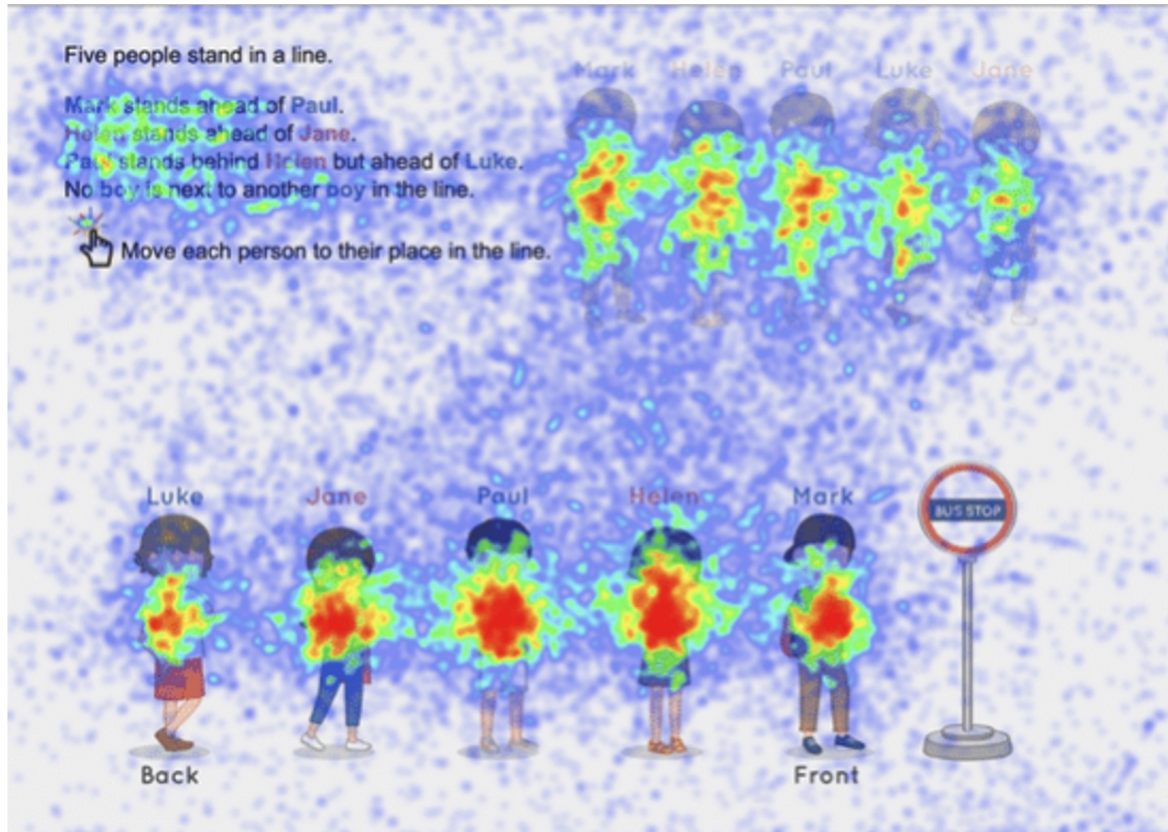


Question Designer

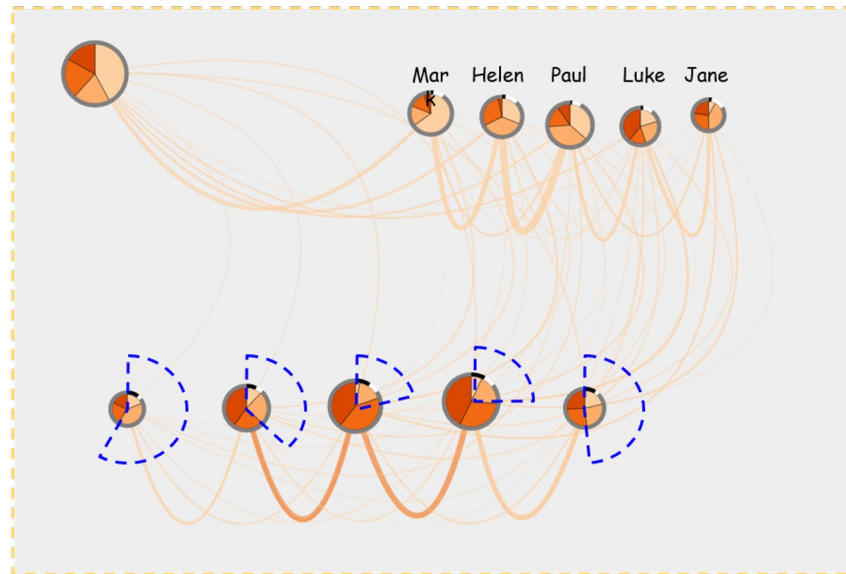
Iterative Design Process - Animation



Iterative Design Process - Heatmap



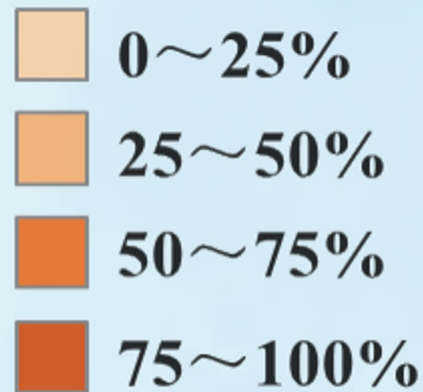
Iterative Design Process - AOI transition graph



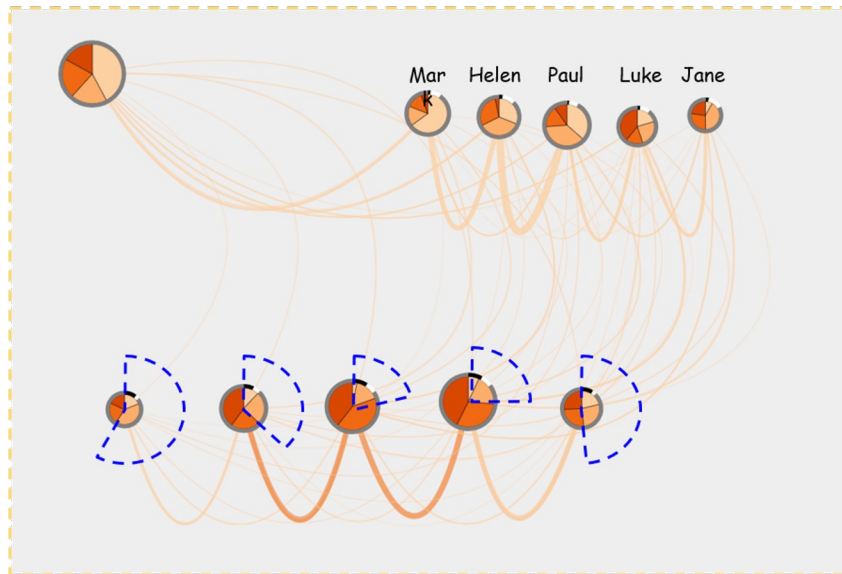
(a)

students with incorrect answers

Time order:

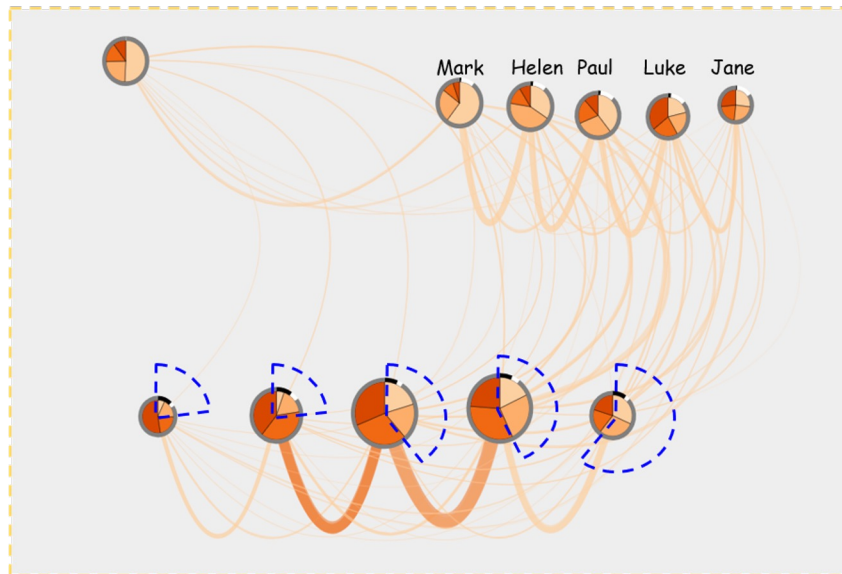


Iterative Design Process - AOI transition graph



(a)

students with incorrect answers



(b)

students with correct answers

1. Data preprocessing: Recover the Steps

For each question:

1

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.

Mark

Helen

Paul

Luke

Jane

2

1

2

3

4

5

6

Luke

Jane

Paul

Helen

Mark

Back

Front

BUS STOP

7

8

9

10

11

For each student:

2 11 4 7 3 8 8 9 ...

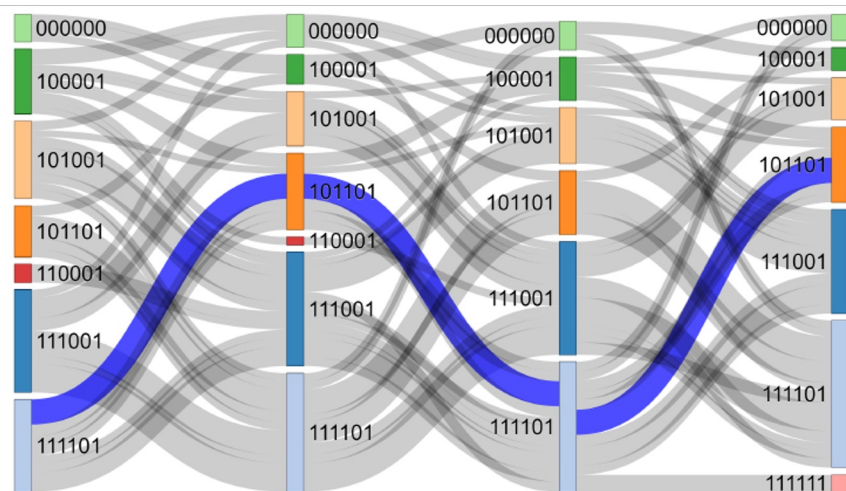
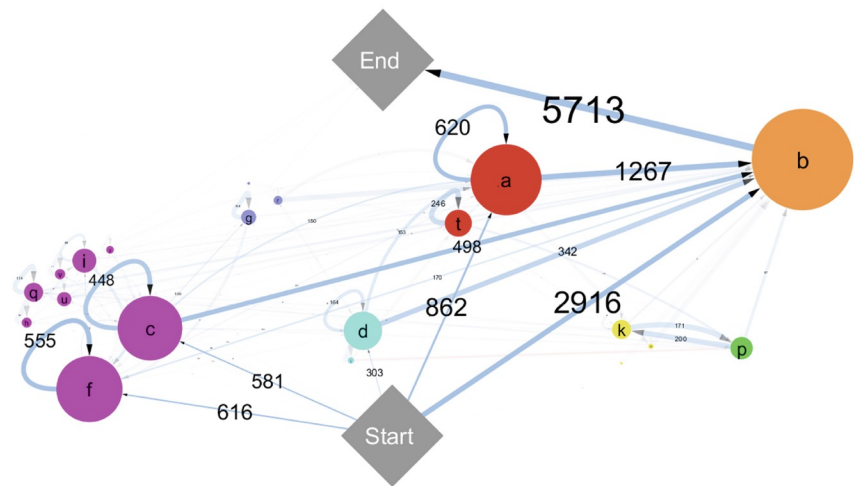
Step1: ,,,Mark

Step2: Paul,,,Mark

Step3: Paul,Helen,,,Mark

Step4: Paul,,Helen,,Mark

...



Not scalable with a very large states.

McBroom, Jessica, et al. "A data-driven method for helping teachers improve feedback in computer programming automated tutors." *International Conference on Artificial Intelligence in Education*. Springer, Cham, 2018.

Wang, Yiting, Walker M. White, and Erik Andersen. "Pathviewer: Visualizing pathways through student data." *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 2017.

2. Data Analysis - State Transition Model

Step: one drag-and-drop

Stage: number of correct conditions

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

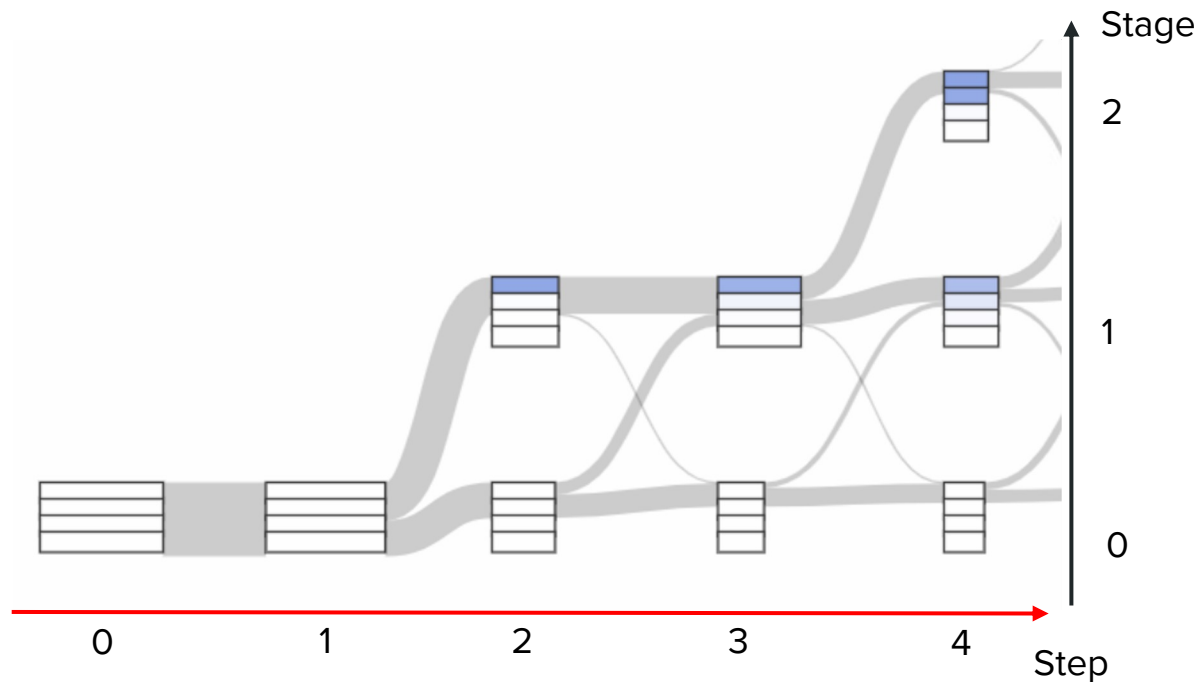
Five people stand in a line.

- Condition 1 Mark stands ahead of Paul.
- Condition 2 Helen stands ahead of Jane.
- Condition 3 Paul stands behind Helen but ahead of Luke.
- Condition 4 No boy is next to another boy in the line.

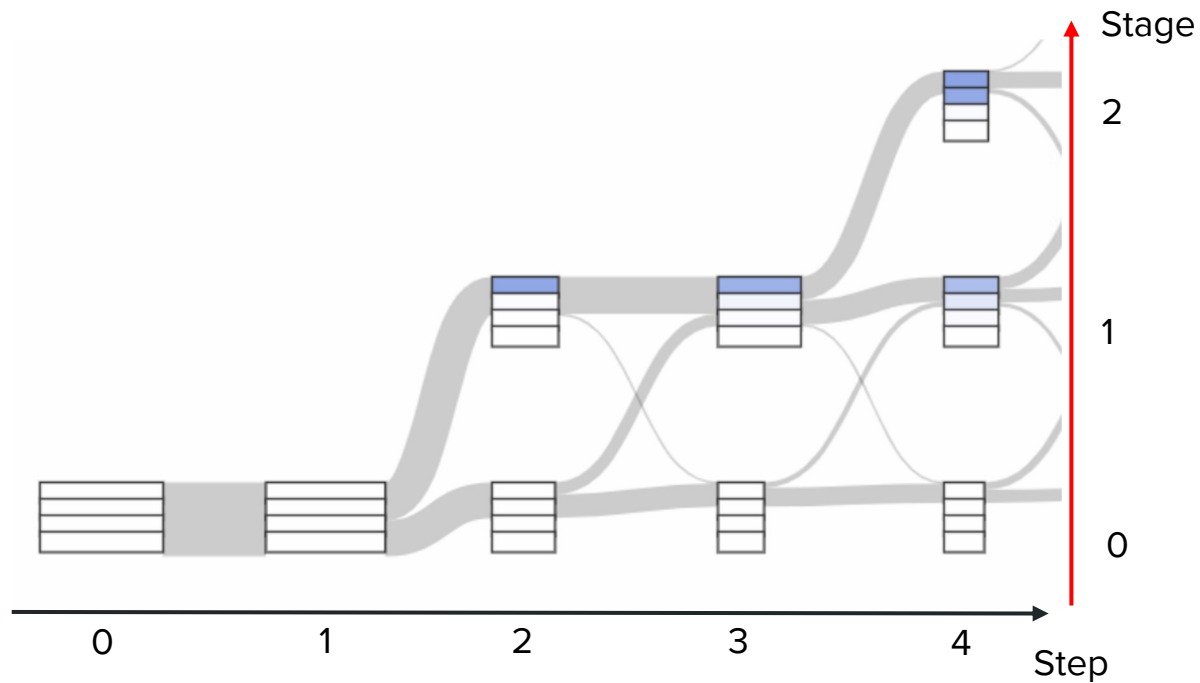


Move each person to their place in the line.

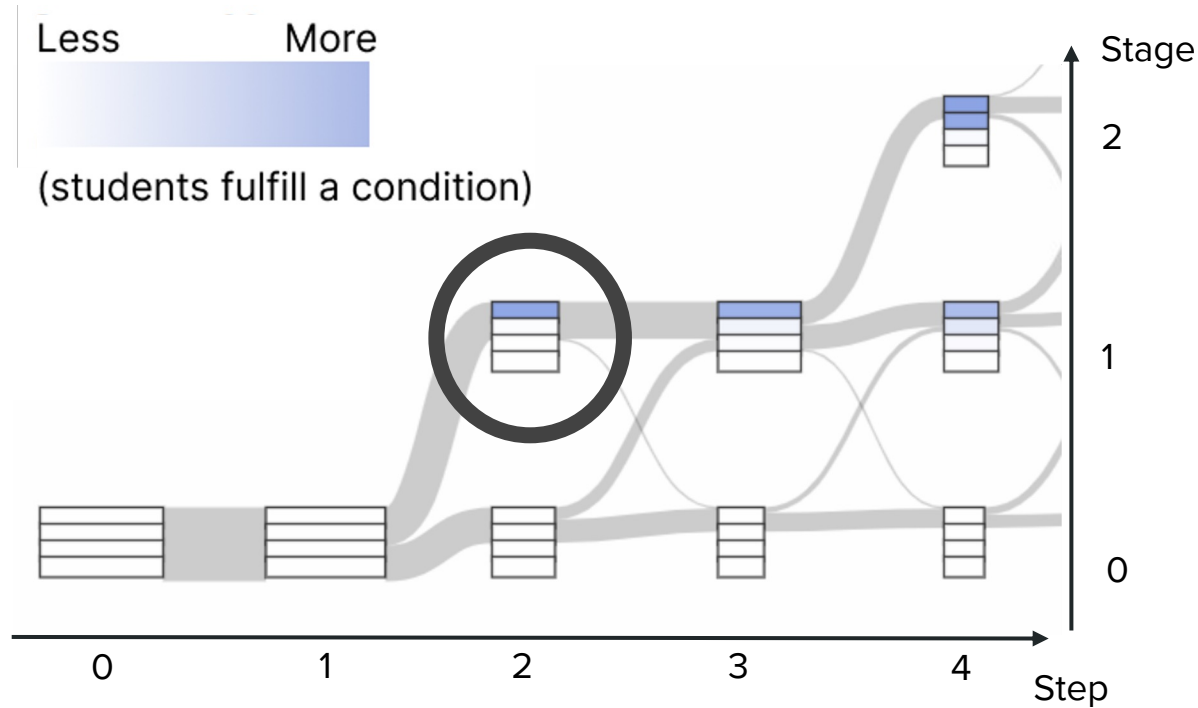
3. Visualization - State Transition Visualization



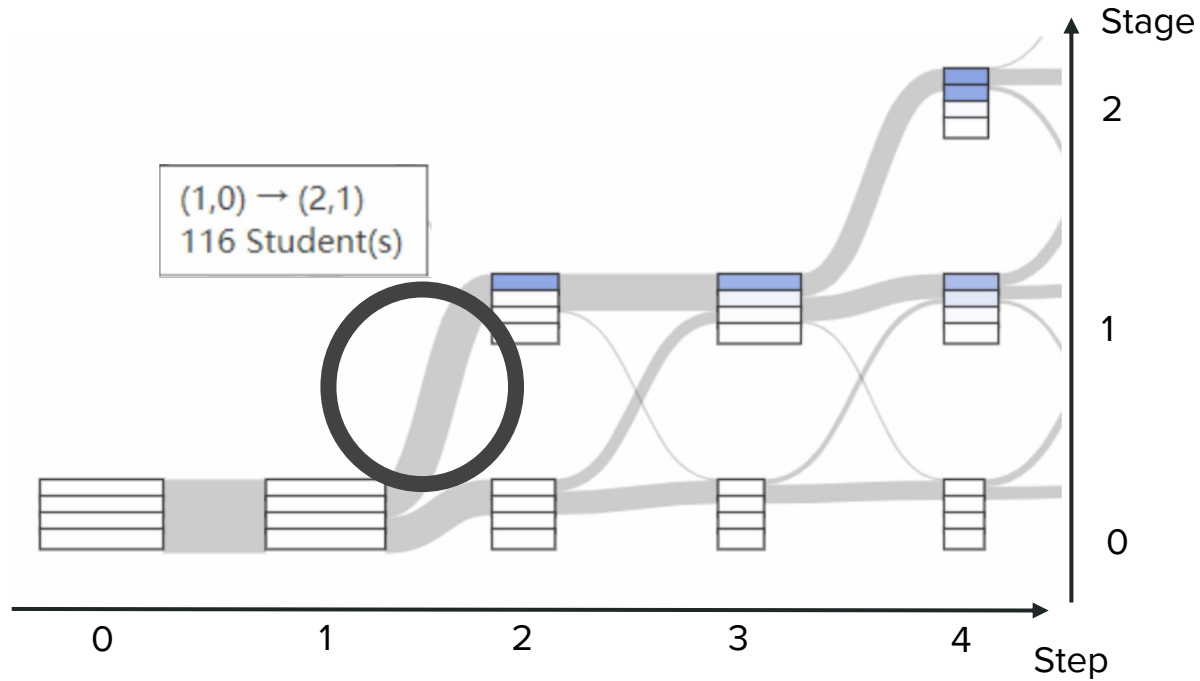
3. Visualization - State Transition Visualization



3. Visualization - State Transition Visualization



3. Visualization - State Transition Visualization



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.

Mark



Helen



Paul



Luke



Jane



Back

Front



Overview

Problem ID: 2 Condition: Relational/Absolute Edit

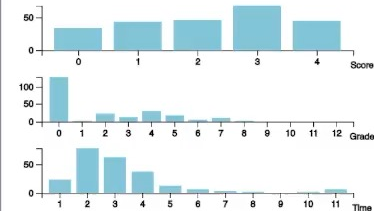
Five people stand in a line.

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.

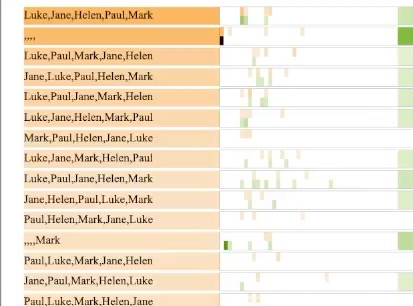
Move each person to their place in the line.



Overall Distribution:



Common Errors:



Transition View

User ID:

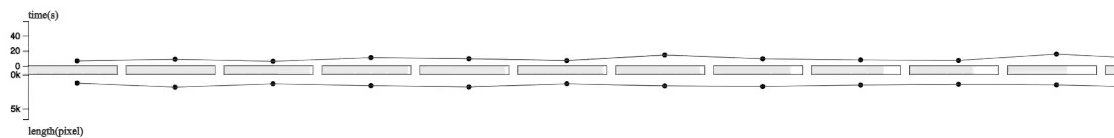
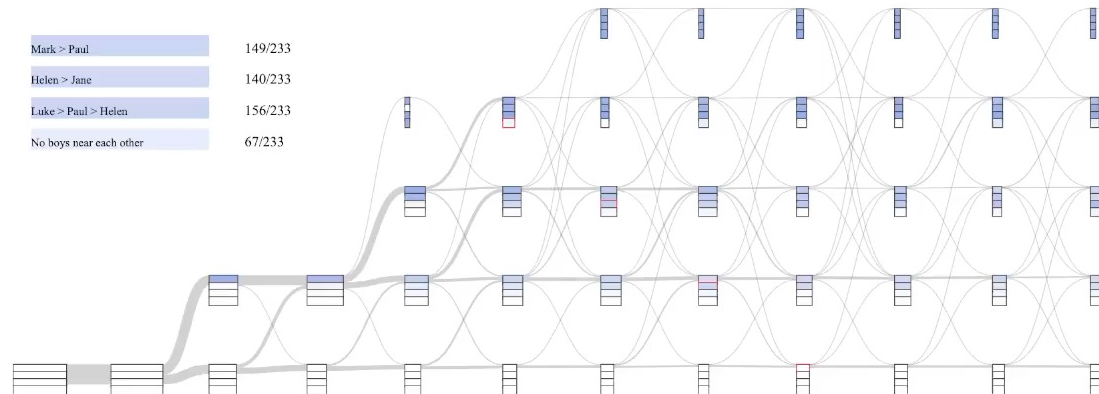
Trans. Freq. Filter: 0

-1

NaviBar:



Mark > Paul	149/233
Helen > Jane	140/233
Luke > Paul > Helen	156/233
No boys near each other	67/233



Comparison View

Less

More

Less

More

(students fulfill a condition)

(students reach a state)

(drops/stops in the level)



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.

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

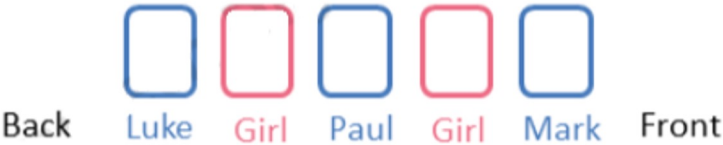
Boy

Girl

Boy

Front

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



"Helen stands ahead of Jane". Therefore,



Overview

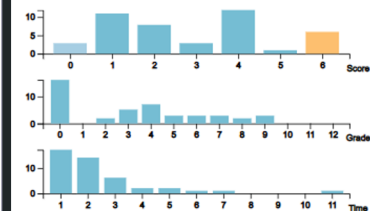
Problem ID: \$ Condition: Relational/Absolute Edit

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

1 2 3 4 5 6

□ □ □ × □ □ □

Overall Distribution:

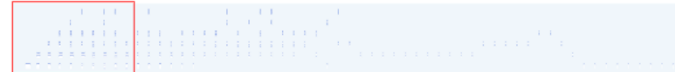


Common Errors:

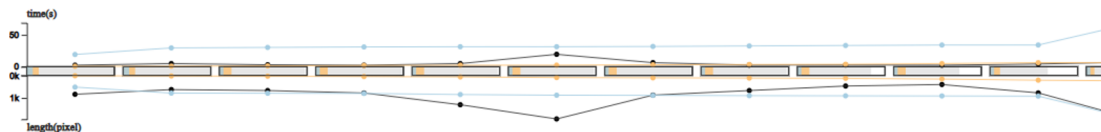
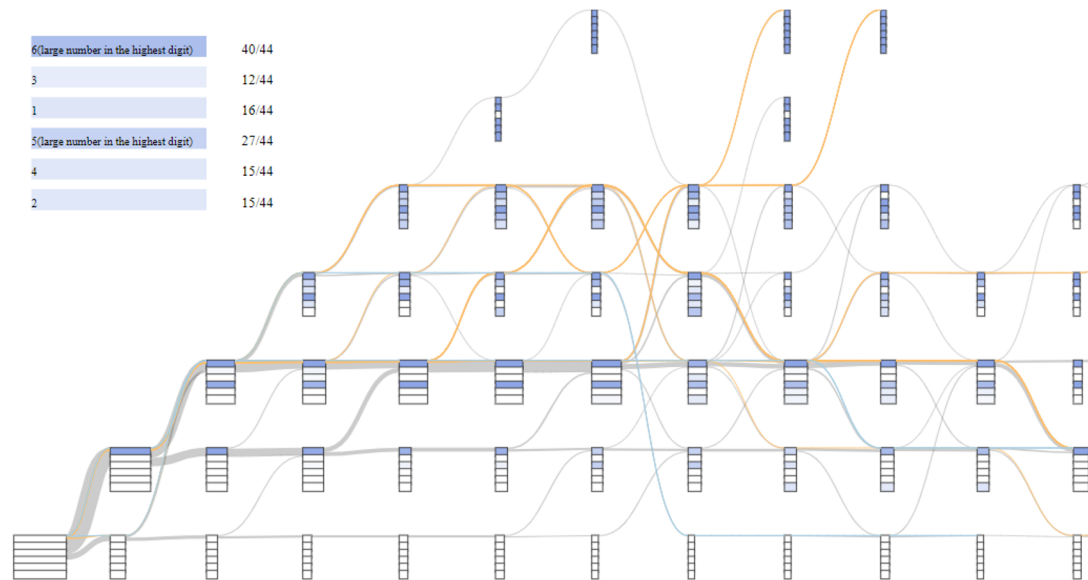
6,5,4,3,2,1	
6,3,2,5,4,1	
6,4,1,5,3,2	
6,4,2,5,3,1	
6,2,1,5,4,3	
.....	
5,3,1,6,4,2	
6,1,2,5,4,3	
6,5,1,4,3,2	
6,4,5,3,2,1	
3,2,6,5,4	
5,2,3,6,1	
5,4,3,2,1	
6,4,2,5,3	
5,.....	

Transition View

NaviBar:



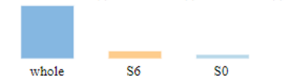
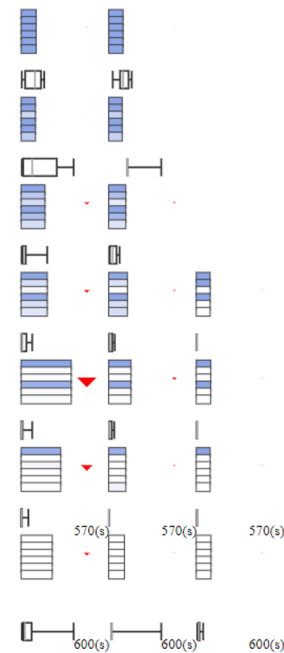
6 (large number in the highest digit)	40/44
3	12/44
1	16/44
5 (large number in the highest digit)	27/44
4	15/44
2	15/44



Comparison View

Less More Less More
(times fulfill a condition) (times reach a state)

(drops/stops in the level)



Overview

Problem ID: Condition: Relational/Absolute Edit

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

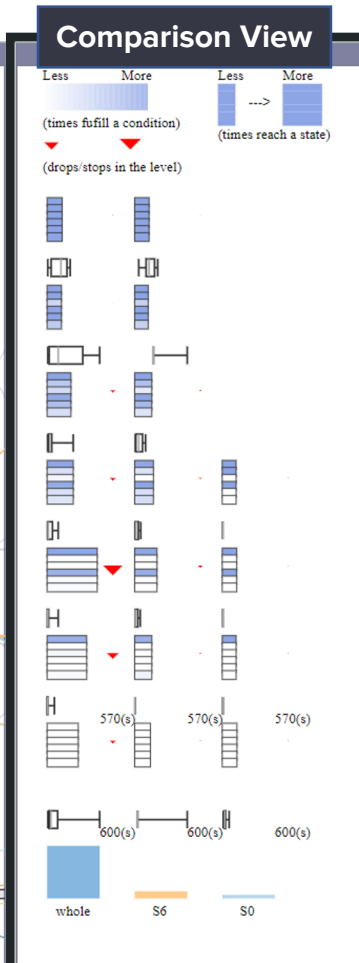
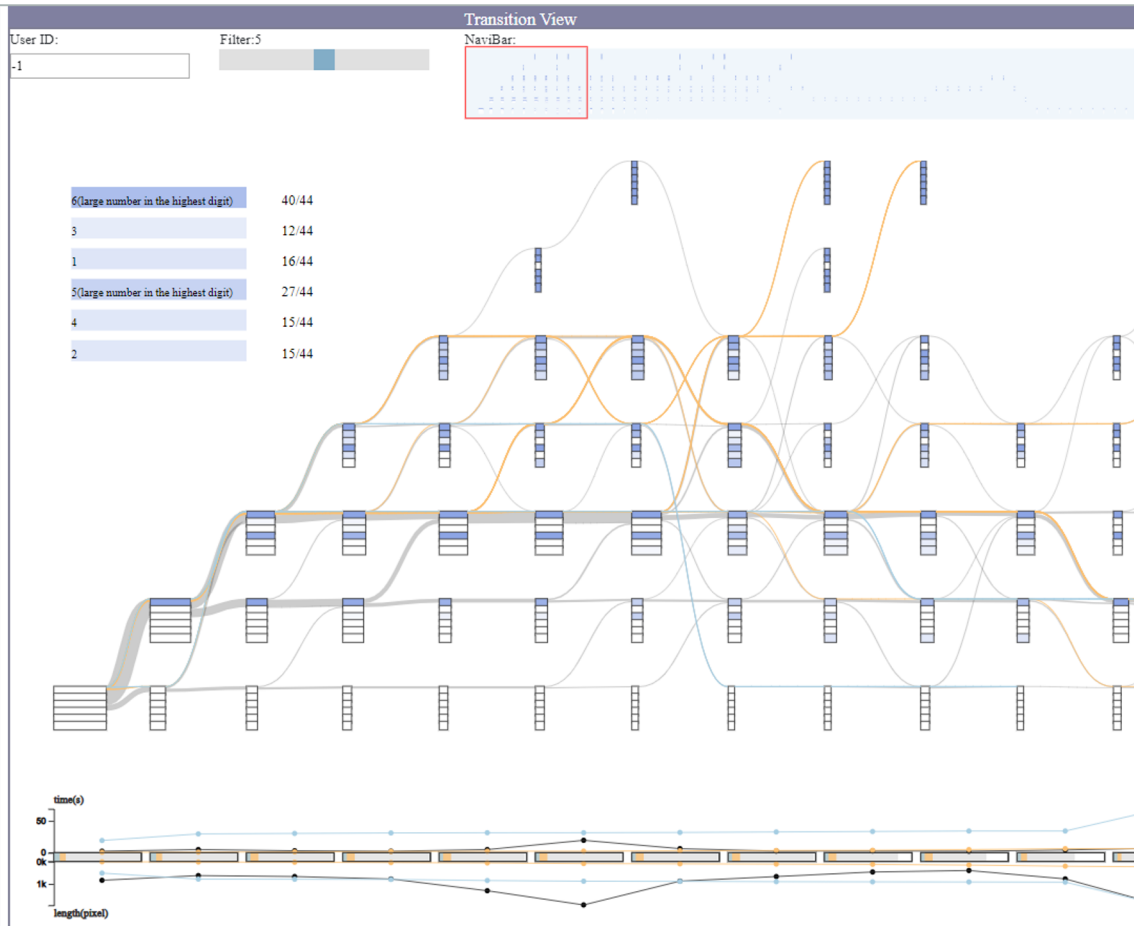
1 2 3 4 5 6

×

Overall Distribution:

Common Errors:

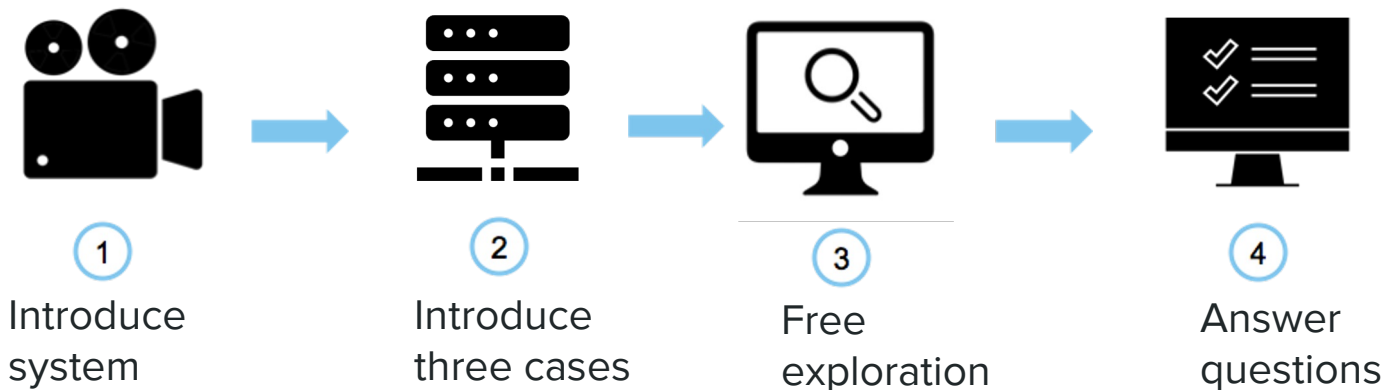
6,5,4,3,2,1	
6,3,2,5,4,1	
6,4,1,5,3,2	
6,4,2,5,3,1	
6,2,1,5,4,3	
.....	
5,3,1,6,4,2	
6,1,2,5,4,3	
6,5,1,4,3,2	
6,4,5,3,2,1	
3,2,6,5,4	
5,2,3,6,1	
5,4,3,2,1	
6,4,2,5,3	
5,.....	



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

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

System usefulness

“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

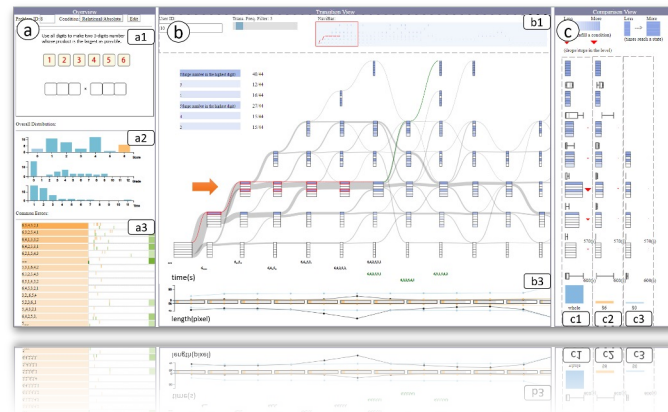
Visual design & interactions

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

--- E7

 Positive
 Neutral
 Negative

QLens visualizes learners' multi-step problem-solving processes to help educators improve learning materials.



Multiple-step problem solving, e.g., how different test cases/rules are satisfied in coding exercises.

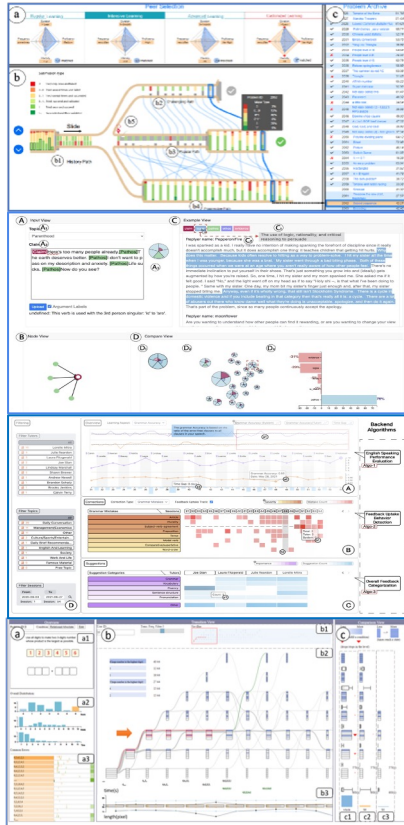


Our collaborator, TrumpTech, uses QLens to improve questions design. The company now serves more than **500 schools** in Hong Kong.

Conclusion

How data can be used for learners and educators in achieving personalized online learning?

Conclusion



Domain situation: formative studies to understand target users' requirements: educators and students

Data/task abstraction:

Data: event sequence data

Tasks: representation, summarization, comparison

Learning behavior modelling:

Learning materials: difficulty level, test knowledge

Students: cognitive skills, non-cognitive variables

Visual encoding: justify alternative designs; address interaction; show the data step by step

Iterative design with educators and students

Lab study, deployment, and post-study interviews



Educators (design loop)



Learners (learning loop)

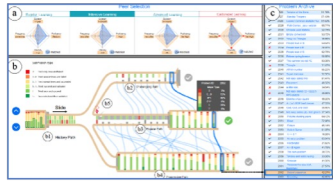
Planning

Performing
Data

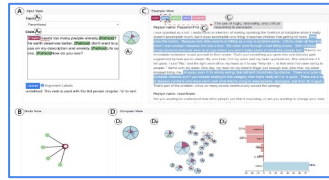
Reflecting

Understanding
and Analyzing

Redesigning



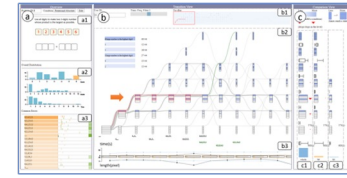
Peerlens (CHI 2019)



Persua (CSCW 2022)



RLens (L@S 2022)



QLens (TVCG 2021)

Enable **learners and educators** to **find insights** from learning data and use the insights to **make decisions** for achieving **personalized online learning**.

Future Research Direction

- More Comprehensive
- More Actionable
- More Effective
- More Context-aware

More Comprehensive

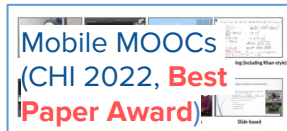
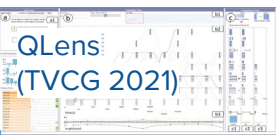
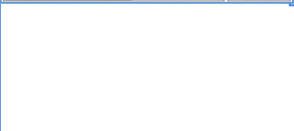
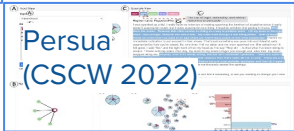


Event Sequence

Text

Code

Video



More Comprehensive - Education data visualization library

Important factors to be considered and explored comprehensively

Target users

e.g. educators, learners

Data Types

e.g. mouse trajectory, event sequence, code, text, video/audio, facial expression, body gestures, performance

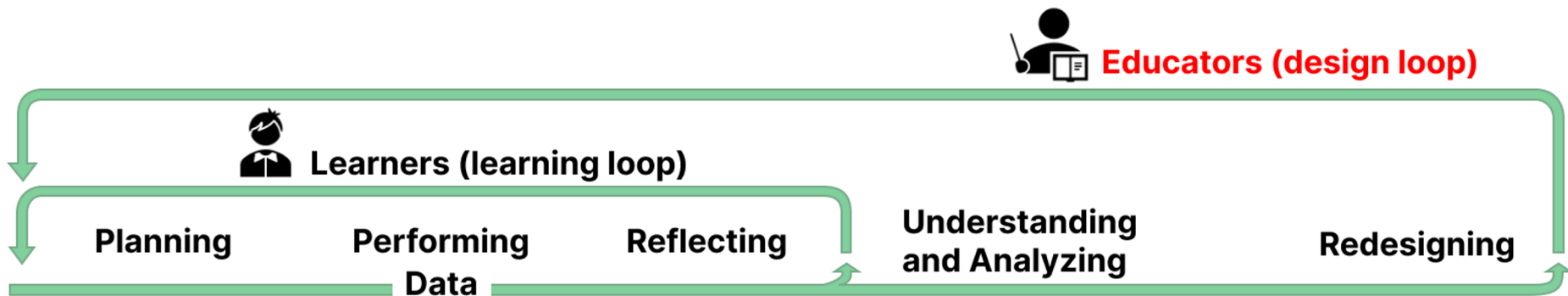
Data Tasks

e.g. summarization, comparison, trend, correlation

Time dimension

Real-time analysis, offline analysis

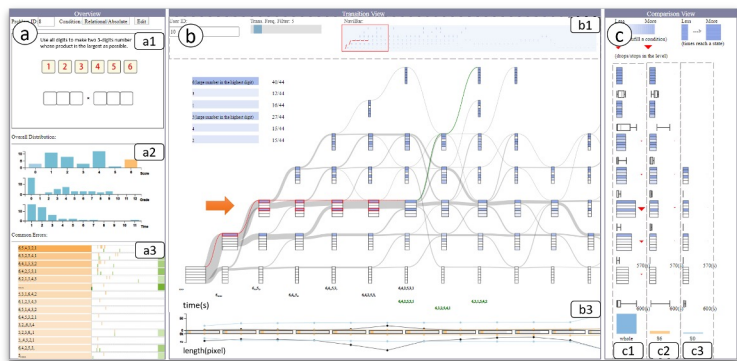
More Actionable



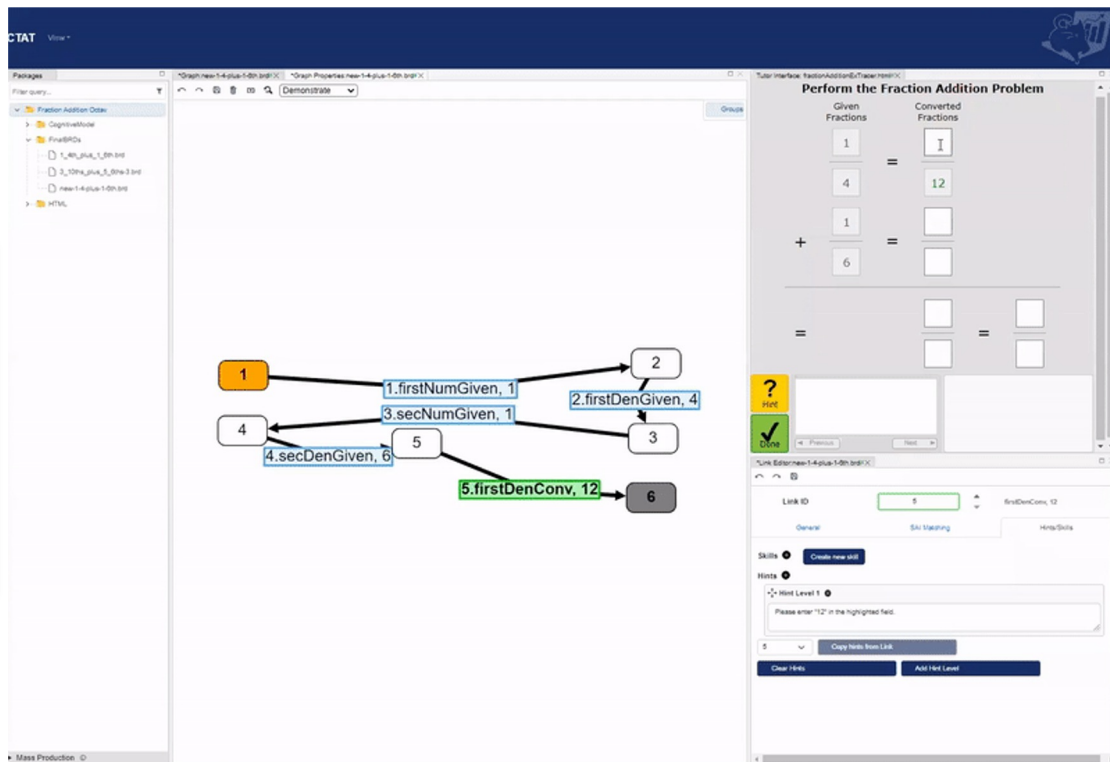
How to push forward learning analytics towards learning design?



More Actionable

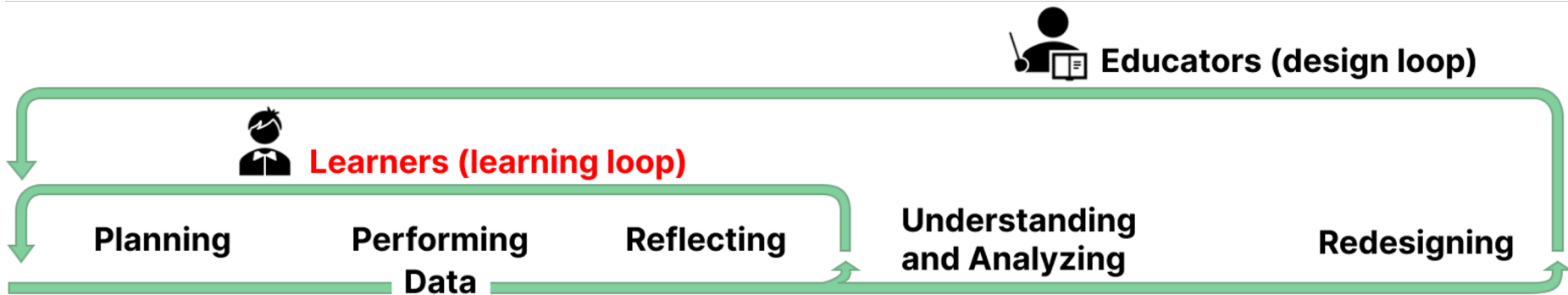


QLens: Learning analytics



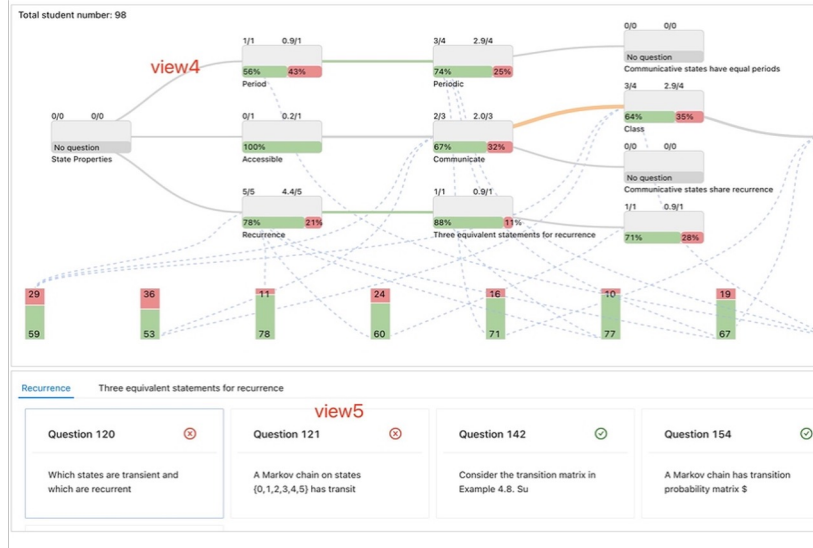
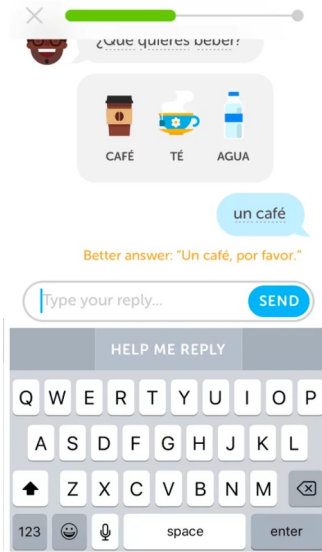
CTAT (cognitive tutor authoring tools):
Learning design

More Effective



*How to infer learners' **accurate** knowledge level and psychology status?*

More Effective

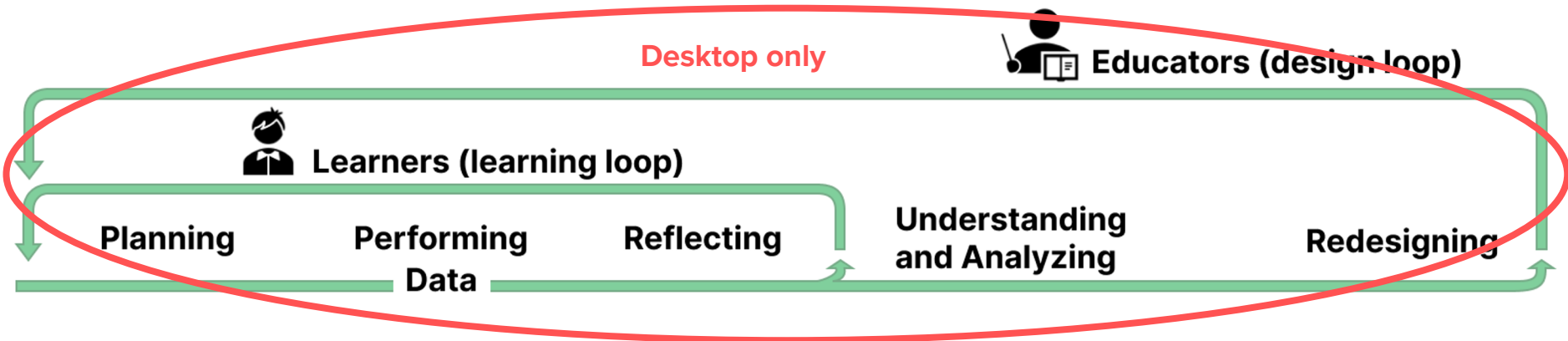


Multiple Rounds Conversations/Interactive Knowledge Map

Other data

Learner-AI collaboration: Reason the Knowledge Level and psychology status

More Context-aware



*What are the **needs**, **challenges**, and **opportunities** of personalized online learning in contexts other than desktop?*

More Context-aware - *Immersive Online Learning*



Learning in VR?

- Cinematography Education on a Soundstage in VR (Xian et al., VR 2023 poster)
- VR Story for Awareness of Covid Spread Threats (ongoing)
- Towards an Understanding of Asymmetric Collaborative Visualization on Problem-solving (Tong et al., VR 2023)

Learning in AR?

- Exploring Interactions with Printed Data Visualizations in Augmented Reality (Tong et al., TVCG 2023, **Honorable Mention Award**)

Other opportunities, e.g., on-the-go learning, digital twin campus?



More Context-aware - *Ubiquitous Online Learning*



Mobile-Friendly Content Design for MOOCs:
Challenges, Requirements, and Design Opportunities
(Kim et al., CHI 2022; **Best Paper Award**)

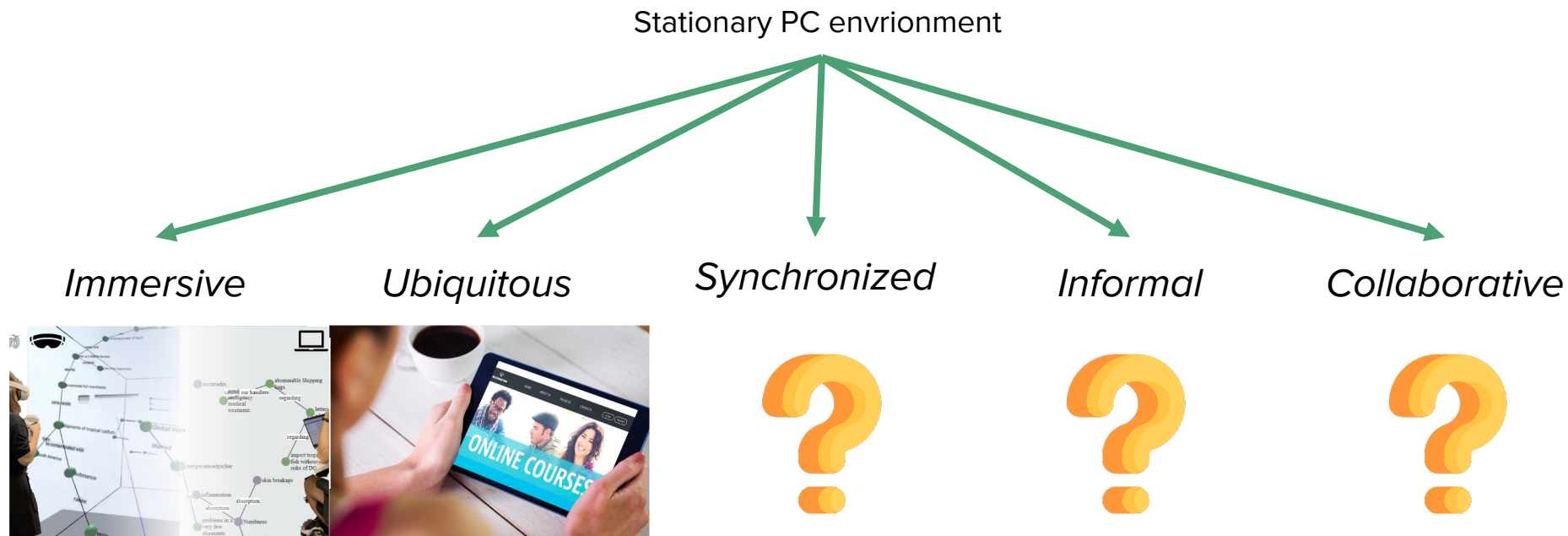


Intelligent tutors on smartphones
(4-year IES project; ongoing)

Interactions on
intelligent tutors?



More Context-aware



- Towards an Understanding of Asymmetric Collaborative Visualization on Problem-solving VR2023

- Exploring Interactions with Printed Data Visualizations in Augmented Reality (Tong, et al., VIS 2022; **Honorable Mention Award**)

Mobile-Friendly Content Design for MOOCs: Challenges, Requirements, and Design Opportunities (Kim et al., CHI 2022; **Best Paper Award**)

More Broad - Enhance Personalization in Other Domains

How about other scenarios like personal health data analysis, investment data analysis, searching engine analysis?



Bias-Aware Design for Informed Decisions

(Zhu, et al., CSCW 2022)

Explaining Air Quality Forecast for Verifying Domain Knowledge using Feature Importance Visualization

(Palaniyappan, et al., IUI 2022)



Welcome to my lab!



**COMPUTER SCIENCE
& ENGINEERING**
TEXAS A&M UNIVERSITY

Join me @Texas A&M



Send me your CV!

mengxia@andrew.cmu.edu



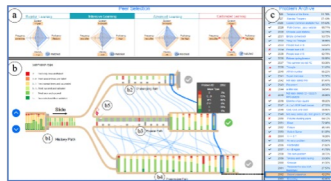
Planning

**Performing
Data**

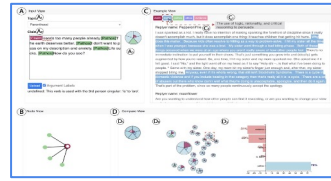
Reflecting

**Understanding
and Analyzing**

Redesigning



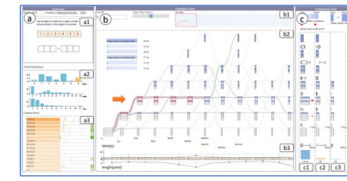
Peerlens (CHI 2019)



Persua (CSCW 2022)



RLens (L@S 2022)



QLens (TVCG 2021)

Human-centered Data-driven Systems for Personalized Online Learning

mengxia@andrew.cmu.edu

Homepage: <https://www.xiameng.org>




Apply here!