

CindyJS によるコンテンツの web 上での操作ログの解析

東邦大学・薬学部 金子 真隆

Masataka Kaneko, Faculty of Pharmaceutical Sciences, Toho University

合同会社三玄舎 中原 敬広

Takahiro Nakahara, Sangensha LLC

東邦大学・理学部 野田 健夫

Takeo Noda, Faculty of Science, Toho University

1 はじめに

一昨年の本共同研究の中で, CindyJS(<https://cindyjs.org/>) により生成されるタンジブルなコンテンツを学習管理システム Moodle 上で提示し, 学習者に PC やスマートフォンなどを用いて操作させるシステムについて紹介した [1]. さらに, 昨年度は記述統計学に関連したコンテンツを学習者に操作してもらった事例を紹介し, グループで操作した際の学習者間の対話や操作プロセスから, 背後にある学習者の思考過程を追跡し, 教授者による適切な介入の方略について知見を得ていくことの重要性について指摘した [2]. 以上をふまえて行った試みの一つが, 大学教養課程の数学で代表的なテーマの一つである, 関数の多項式近似に関する事例である [3][4][5]. これまで大学 1 年生に該当の内容を講義してきた経験からすると, 多項式近似の計算公式を記憶し, 与えられた関数にこれを適用して近似式を計算する作業については大方の学生がこなせている反面, この計算による近似プロセスがイメージできているか, さらに, 付随して無限小の次数の概念がつかめているかということになると, 甚だ心許ないというのが実態であった. これらの点はいずれも, 数式による提示のみで一般的な学生に理解させるのが容易でなく, テクノロジーを用いたアプローチに一定の効果が期待できる部分である [6]. 試行にあたっては図 1 にあるコンテンツを用意し, $x = 0$ の周りで関数 $y = \sqrt{1+x}$ の 3 次近似式

$$\sqrt{1+x} \sim a + bx + cx^2 + dx^3$$

を探索させた.



図 1 使用した CindyJS によるコンテンツ

図1において、左側に配置された4つのスライダーを指で操作することにより、係数 a, b, c, d を動かすことができ、対応した3次関数のグラフが即時に描画される仕様になっている。この場合、1次以下の項 $a + bx$ が接線の方程式を与える形 $f(0) + f'(0)x$ と異なっていると、2次以上の係数をどのように調整しても $x = 0$ の周りでの近似が良くなり、このような状況を観察することによって、近似の剰余項 $o(x^2)$ が2次の無限小となる状況が理解されることが期待されている。さらに、操作の中で学習者が直面すると想定される課題は以下の2点となる。

1. そもそも近似はどの範囲で可能なのか？
2. $\frac{f''(0)}{2}x^2, \frac{f'''(0)}{6}x^3$ はどのような「調整」をするための項か？

関数 $y = \sqrt{1+x}$ の場合、 $x = -1$ に特異点を持つことから、 $x = 0$ を中心とした多項式近似の収束半径は1であり、次数が高くなればなるほど良い近似が得られることになるが、近似を3次までに限定した場合、 $x = 0$ を中心としたどこまでの範囲で近似を良くするかによって、最適な係数の取り方が変わってくることになり、これが課題1の背景である。課題2に関する疑問が出てきた場合には、点 $(0, 1)$ 付近をズームアップした形のコンテンツが必要になり、実際そのような提示も行っているが、本稿では割愛する。

2 教育工学的背景

本稿で紹介する事例のように、数学的な事実が教授者から学習者に対して伝達されるのではなく、具体例を用いた探索活動を通して理解を深めていくタイプの学びを、教育工学では inquiry based learning と呼んでいる。一口に探索活動と言っても、いろいろなアプローチが有り得るが、Piaget や Papert に連なる constructivistic なアプローチをとった場合に、効果を最も発揮しやすいことがしばしば指摘される [7]。これをもう少し敷衍すれば、学習者を ill-structured かつ open-ended な問題に取り組みさせることが望ましいということになるであろう [8]。上述の通り、本稿で紹介する事例もこのような線に沿ったものと位置づけられるが、このようなアプローチは教育効果に対する多くの期待がある反面、その検証が必ずしも易しくないとの指摘が少なくないことに注意すべきである。実際、モバイルデバイスなどのテクノロジーを用いた教育の効果に関する研究をレビューしたメタ解析の論文の多くで、こうした研究では効果検証に用いるスケールとテクノロジーを用いている際のプロセスとの関係について追究する必要があったにもかかわらず、プロセスログの分析が決定的に欠けており、これによって研究ごとの結果のばらつきが生じてしまっている点が指摘されている [9]。

教育工学の立場からこのような学習活動をとらえる際、多くの場合に分析の枠組みの基盤となるのは Vygotsky による cultural-historical approach である。Vygotsky はその著書の中で、「目に見える対話や探索活動は、social context の中に位置づけられることで学習者により内部化 (internalize) される。学習者が複雑な問題に取り組む際には “zone of proximal development” が生じるので、さまざまな補助が必要である」と主張している [10]。このさまざまな補助になりうるものの一つがテクノロジーだということになる。

こうした Vygotsky の考え方を現代に受け継ぎ、テクノロジーの教育利用に関する研究の主流をなすのが computer supported collaborative learning (CSCL) のパラダイムであるが、その始祖というべき Koschmann は、研究手法として依拠する ethnomethodology における考え方 [11] を引用しつつ、CSCL に属する教育手法は “tutorial problem” の一種と見なしうるものであって、その本質は、自らが行っていることが何物であるのか、学習者に認識させるためのガイドとなる点にあることを主張している [12]。この主張は、一見すると学習者のみを対象としたもののようにも見えるのだが、CSCL における学習活動が tutorial problem の枠組みの中で行われることを踏まえると、CSCL の研究者が学習者の思考プロセスを理解する上で、その行動のひとつひとつを丁寧に意味付けていくという、ethnomethodological なアプローチが不可欠であることを同時に主張している点に注意が必要だと考えられる。

CSCL が包摂しうる学習活動は多岐にわたるが、本稿で紹介する事例が分析上の枠組みとして依拠しうるのは、Kapur らによって主唱された productive failure instructional design であろう [13]。CSCL に基づく教育実践が増加するにつれ、学習者を補助するツールがいろいろと開発されてきたが、学習者の助けとなることが期待される一方で、argumentation tool や scripting interaction などの利用によって学習者の思考が一方向に lock-in されてしまい、問題解決に向けた十分な追究をかえって制限してしまいかねない危険性も同時に指摘されてきた [14]。Kapur らは、本稿におけるような ill-structured problem に学習者が取り組むにあたり、外部からの補助を行わないことで、短期的には学習の失敗をもたらすように見える一方、長期的には隠れた効率性 (hidden efficacy) をもたらしうることに注目したのである。こうした場合、学習プロセスは著しく複雑化するのが常であり [15]、プロセスデータの取得・分析に関して一段とハードルが高くなることを覚悟しなくてはならない。

3 事例研究の方法と結果

CindyJS によるコンテンツは html 形式にフォーマットされているため、オンラインの条件下であれば一般的なウェブブラウザのみで動作させることができるが、オフラインの状況であっても、必要な「ライブラリ」を取り込んでおきさえすれば、アプリケーションのインストールなどを要せずに同様の動作を確保できる。特にまとまった人数で操作させる場合、ネットワークの安定性に十分な信頼をおけないケースがほとんどであるため、授業時の試行の場合は iPad にライブラリとコンテンツを装備した上で、オフラインの状態ですべて操作させることがほとんどであった。本節で紹介する試行も同様の条件で行っている。数年前からようやく iPad 上での画面録画が可能になってきていることを踏まえ、最近実施している試行では、学習者による操作画面をすべて録画し、これを行動分析ソフトとして広く利用されている Sportscode (<https://www.hudl.com/elite/sportscode>) に取り込んだ上で、操作プロセスの時系列的なコーディングを行ってきている。大学1年生35名(男性のみ)にひとり1台ずつiPadを配布し、コンテンツを操作してもらった際は、事前に講義の中で多項式近似の公式の導出や無限小の次数の概念に関する数式的な説明を与えた上、いくつかの関数について、各次数ごとに近似した結果のグラフをス

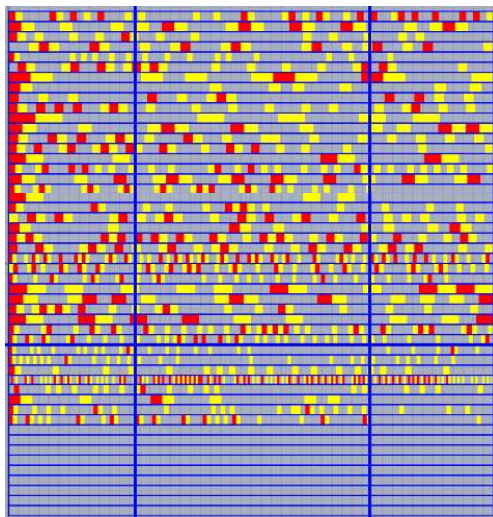


図4 全被験者の操作プロセスの可視化

図4から明らかなように、操作パターンは被験者によって大きく異なっており、探索上の方略やその背後にある思考過程について、被験者の間に大きなばらつきが存在したことを示している。一方で、操作プロセスを細分した上で低次の係数 a, b の操作比率を追跡してみると、プロセスの終盤に向けて低下している事例が多いことが観察される。同一のコンテンツをグループに操作してもらい、その操作プロセスとメンバー間の対話との連関について調べた筆者の先行研究 [4] によると、低次の係数の操作比率の低下が、大域的近似から局所的近似へと学習者の意識が遷移している状況を示すものであることが強く示唆されている。この示唆を前提とした場合、図4に示される上述のような操作ログデータの傾向の遷移から、学習者の思考プロセスについて何が読み取れるだろうか。唯一つ考え得るのは、操作の途中段階で大域的な近似を探すべく低次の係数を動かしてみたものの、1次近似の部分が接線を与える形と異なる場合には高次の係数をどのように調整しても良い結果が得られないという体験を通じ、高次の無限小の概念を直感的に把握して行っているプロセスである。これはいわば失敗を通して学びを深めていく過程と見なすことが可能であり、以上のような学習者の探索活動が *productive failure* の一種と見なせるという、本稿の主張の根拠をなしている。

では、このような見方を前提とした場合に、図4に示される大学生による操作と高校生による操作との比較はどのような示唆をもたらすだろうか。特に高校生の分のサンプル数が十分でないため、確定的なことは言えないが、操作プロセスの後半部分において、大学生の群と比べ、高校生の群で低次の係数の操作比率の低下が顕著であることが観察できる。これは、特に高校生の群で、操作プロセスの前半において上述のような体験的な学習がより深化したことを示唆するものである。以上の結果は、同一のコンテンツを学習者が操作するとしても、教授者がどのように介入・補助することが適切か、学習者のプロファイルに応じて相当慎重に判断しなくてはならないことを鮮明に示している。今後、このような結論の妥当性について検討するため、さらに多くの高校生を被験者として事例研究を継続する予定である。

4 総合論議と今後の研究の方向性

センシングや通信に関わる技術の進歩に伴って、最近では学習プロセスに関する粒度の高いデータが取得できるようになってきており、一部には生体信号を用いた multimodal な分析を行う研究も見られる [18]. こうした研究からは、学習の背後で動いている学習者の感性やメタ認知に関する知見が得られるようになってきている一方で、本稿で扱うような認知的な側面について知見を得ようとすると、学習者の発話を含めた質的なプロセス分析を絡めざるを得ないのが実情である. そうした場合、カテゴリ付けされた発話の頻度の推移といったテキストマイニングの手法を用いるのが一般的であり、筆者の先行研究 [3][4] もその方向性に従ったものであるが、分析手法の性格上、一定の時間幅をとって発話情報を集積することが前提となる. 結果的に、その時間幅に対して思考プロセスの性格付けを行える可能性は高いのだが、本稿の事例のように時々刻々と内容が変化していくような場合には、データとしての粒度が確保できず、十分な分析に至らない危険性があることが、最近のいくつかの研究で指摘されている [19]. これを補うもののひとつと位置づけられるのが、発話の発生順序に着目した sequential analysis であり、最近では発話の sequential pattern の傾向と学習成果・潜在能力との関連性を調べた研究も出てきている [20]. 発話と操作との違いはあるものの、前節で紹介した事例もこのような sequential analysis の一種であり、そこで観察された結果は [20] と同様の知見をもたらす得ることが示されているとも言えるであろう.

しかし、前節で得られているパターンデータでは、依然として学習者の操作に含まれる多くの時系列情報が失われてしまっていることに注意が必要である. プロットされているのはあくまでどの係数をどの順番に操作したかという情報だけであり、係数を動かした範囲や操作時間にかかわる情報は含まれていない. 学習者がある時点で行った操作の結果を観察し、その結果をそれ以後の操作に反映させていることを考えると、これでは操作プロセス全体を解析する材料として十分とは言えないことが明らかであろう. 学習者の対話についても、過去の発話との深い関連性がプロセス全体を著しく複雑化させるのが常であり [21], こうした場合、Hidden Markov Model に代表されるような sequential analysis の枠組みでは必ずしも十分とは言えず、よりイベントの中身に密着した分析が求められることが指摘されている [22].

以上のような操作ログデータの分析に関する問題を解決するには、ethnomethodology の発想と同様に、まずは操作の詳細をありのままに記録することが求められると考えられる. そこで第2著者(中原)は、学習者がコンテンツの点要素を操作した経過をタイムスタンプつきで集積して csv 形式にフォーマットし、操作終了後のクリック動作のみでダウンロードできる Moodle プラグインを新たに開発した. 図5は、そのプラグインを備えた Moodle のコースのスクリーンショットである. 基本的に、このページにアクセスした瞬間からログの集積が開始され、操作終了後にページを閉じると自動的に集積を終了する仕組みになっている. 学習者個々に異なる ID で Moodle にログインする形になるため、同時に複数の学習者が操作をした場合でも、操作者個々に異なるログとして蓄積されることになる. ページの右上にあるメニューから「Cindy Log Report」というタグを選択すると、集積されているデータの一覧が表示されるので、必要なものを選んでクリックすると、やはり自動的にダウンロードが開始される.

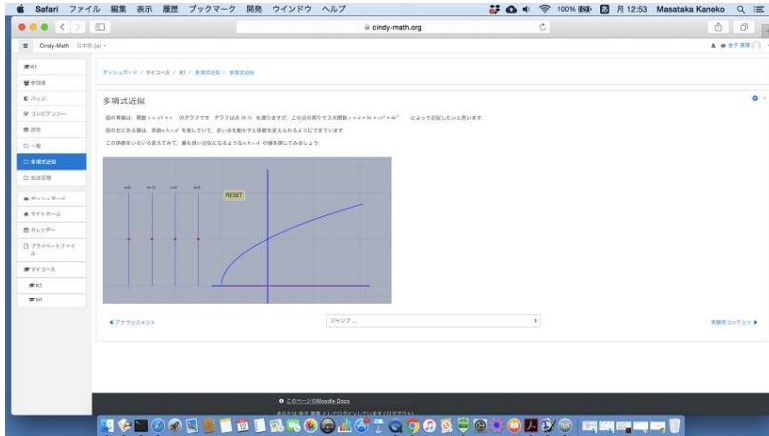


図5 学習者がコンテンツを操作するページの画面

このプラグインが出来てからまだ日が浅く、本稿のコンテンツを使った試行もわずかに一例にとどまっているが、図6はこの事例について、csv ファイルをダウンロードした際の画面である。学習者による操作が行われる度に、操作時刻、操作した点、および点の移動先の座標が記録されるようになっている。

ID	Download ID	Action	Target	X	Y	Timestamp	
1	4	3	19:27:48	K	-1.020491	1.592E+2	37343
2	4	3	19:28:00	K	-1.020594	1.592E+2	37342
3	4	3	19:28:00	K	-1.020677	1.592E+2	37338
4	4	3	19:28:00	K	-1.020736	1.592E+2	37336
5	4	3	19:28:00	K	-1.020832	1.592E+2	37412
6	4	3	19:28:00	K	-1.020932	1.592E+2	37418
7	4	3	19:28:00	K	-1.021138	1.592E+2	37444
8	4	3	19:28:00	K	-1.021353	1.592E+2	37461
9	4	3	19:28:00	K	-1.021466	1.592E+2	37478
10	4	3	19:28:00	K	-1.021576	1.592E+2	37485
11	4	3	19:28:00	K	-1.021692	1.592E+2	37511
12	4	3	19:28:00	K	-1.021808	1.592E+2	37528
13	4	3	19:28:00	K	-1.021926	1.592E+2	37545
14	4	3	19:28:00	K	-1.022045	1.592E+2	37579
15	4	3	19:28:00	K	-1.022165	1.592E+2	37591
16	4	3	19:28:00	K	-1.022286	1.592E+2	37608
17	4	3	19:28:00	K	-1.022408	1.592E+2	37625
18	4	3	19:28:00	K	-1.022531	1.592E+2	37642
19	4	3	19:28:00	K	-1.022654	1.592E+2	37659
20	4	3	19:28:00	K	-1.022777	1.592E+2	37676
21	4	3	19:28:00	K	-1.022901	1.592E+2	37693
22	4	3	19:28:00	K	-1.023025	1.592E+2	37710
23	4	3	19:28:00	K	-1.023149	1.592E+2	37727
24	4	3	19:28:00	K	-1.023273	1.592E+2	37744
25	4	3	19:28:00	K	-1.023397	1.592E+2	37761
26	4	3	19:28:00	K	-1.023521	1.592E+2	37778
27	4	3	19:28:00	K	-1.023645	1.592E+2	37795
28	4	3	19:28:00	K	-1.023769	1.592E+2	37812

図6 操作ログデータファイルのサンプル

当然のことながら、この事例のように点の移動が頻繁に行われるような場合、データ量は非常に大きくなり、取り扱いはずしも容易ではない。しかし、操作プロセスの素朴な観察等から、学習者の操作方略に付随した特徴的な操作パターンが予測された場合、得られたアウトプットを加工して、そのパターンの時系列的な推移を調べられる可能性が生じる。したがって、うまくいけば、操作パターンのシグナルに基づき教授者がいつどのように介入したらよいかという点について、一定の指針が得られる可能性がある。たとえば本事例の場合、前述のとおり操作パターンが近似を目指す範囲の大きさを示すシグナルである可能性が示唆されているが、図6のログデータを加工すれば、良く近似できている範囲の大きさを示す数値を計算し、その時系列的な推移を追跡できる可能性がある。近似範囲を表す指標としてはいくつかのものが考えられるが、ここでは ε -収束

半径と呼ぶ量を計算する。これは、事前に設けた一定の閾値 ε に対し

$$\sup_{-c \leq x \leq c} |\sqrt{1+x} - a - bx - cx^2 - dx^3| \leq \varepsilon$$

を満たすような c の上限で与えられる数値である。図7は、4つの ε の値 (0.1, 0.2, 0.3, 0.4) に対し、JAVA でプログラムしたマクロを用いて図6のログデータの各時刻で ε -収束半径を計算し、csv形式にフォーマットしたものの様子である。

時刻	ε=0.1	ε=0.2	ε=0.3	ε=0.4
1	0	0	0	0
2	0.000000	0.000000	0.000000	0.000000
3	0.000000	0.000000	0.000000	0.000000
4	0.000000	0.000000	0.000000	0.000000
5	0.000000	0.000000	0.000000	0.000000
6	0.000000	0.000000	0.000000	0.000000
7	0.000000	0.000000	0.000000	0.000000
8	0.000000	0.000000	0.000000	0.000000
9	0.000000	0.000000	0.000000	0.000000
10	0.000000	0.000000	0.000000	0.000000
11	0.000000	0.000000	0.000000	0.000000
12	0.000000	0.000000	0.000000	0.000000
13	0.000000	0.000000	0.000000	0.000000
14	0.000000	0.000000	0.000000	0.000000
15	0.000000	0.000000	0.000000	0.000000
16	0.000000	0.000000	0.000000	0.000000
17	0.000000	0.000000	0.000000	0.000000
18	0.000000	0.000000	0.000000	0.000000
19	0.000000	0.000000	0.000000	0.000000
20	0.000000	0.000000	0.000000	0.000000
21	0.000000	0.000000	0.000000	0.000000
22	0.000000	0.000000	0.000000	0.000000
23	0.000000	0.000000	0.000000	0.000000
24	0.000000	0.000000	0.000000	0.000000
25	0.000000	0.000000	0.000000	0.000000
26	0.000000	0.000000	0.000000	0.000000

図7 ε -収束半径の計算結果

さらに図8では、この結果をもとに、 ε -収束半径の時間推移を可視化している。

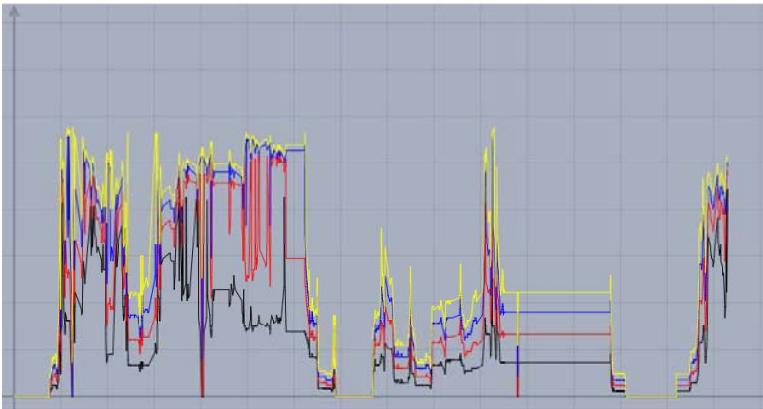


図8 ε -収束半径の時間推移

図8を見ると、操作プロセスの比較的早い段階から、 ε -収束半径が大きい状態が続いており、最適解に近い近似が早期に得られていたことを示している。これまでの事例研究で、このように早期に良い近似が得られ(てしまっ)たケースでは、逆に十分な探索が尽くされず、ごく限られたケースを見ただけで探索を終えてしまうケースが少なからず確認されていた。操作状況を観察しながら、この試行でも同様の事態の発生が懸念されたため、操作途中で2度の介入を行った。これは図8の中で、 ε -収束半径が0になっている2つの時間帯に該当する。最初の介入では、より大域的な近似を目指してみるよう

にとの指示を与えたが、想定通り近似が良くならず、探索プロセスに乱れが生じていたことが観察される。2度目の介入では、手前の操作結果に関する振り返りを促し、結果として再び開始当初の状態に戻っていることも確認できる。

このような様子を見ると、Mercer[21]が指摘するCSCLにおける学習者の対話の複雑さと同様に、動的コンテンツの操作の場合も、各時点までのプロセスに関するわずかなゆらぎがそれ以後の進行を大きく左右し、結果的に高度に複雑なログデータを発生させる状況が明らかである。荒唐無稽との誹りを恐れずに言うならば、動的コンテンツの操作はある意味で学習者と数理モデルの対話とみなすことができ、したがって、そのプロセスの分析には対話分析の場合と同様に粒度の高いデータが必要になるということができよう。本稿で紹介した新しいシステムは、そのようなデータを手軽に取得できる有力な手段になるものと期待される。

謝辞

本研究はJSPS 科研費 18K02872, 19K03175, 平成30年度東邦大学理学部教養科大型予算の助成を受けている。

参考文献

- [1] 金子真隆, 中原敬広, 中村泰之: CindyJS によるコンテンツの Web 上での利用について, 京都大学数理解析研究所講究録 2067, pp.37-46, 2017
- [2] M. Kaneko, T. Noda: Learning descriptive statistics with the use of dynamic content generated by CindyJS, 京都大学数理解析研究所講究録 2105, pp.137-145, 2018
- [3] 金子真隆: CindyJS によるコンテンツを用いた協調学習における操作と対話との関連の追跡 (I), 日本教育工学会第 34 回全国大会講演論文集, pp.279-280, 2018
- [4] 金子真隆, 野田健夫: CindyJS によるコンテンツを用いた協調学習における操作と対話との関連の追跡 (II), 日本教育工学会研究報告集, 18-5, pp.45-52, 2018
- [5] 金子真隆, 江木啓訓, 中原敬広, 野田健夫: CindyJS によるコンテンツを用いた協調学習における操作と対話との関連の追跡 (III), 日本教育工学会 2019 年度秋季全国大会講演論文集, pp.289-290, 2019
- [6] J. Bookman, D. Malone: The nature of learning in interactive technological environments: A proposal for a research agenda based on grounded theory, *Research in Collegiate Mathematics Education*, 5, pp.182-204, 2003
- [7] Q. Li, X. Ma: A meta-analysis of the effects of computer technology on school students' mathematics learning, *Educational Psychology Review*, 22, pp.215-243, 2010
- [8] C. E. Hmelo-Silver: Problem-based learning: What and how do students learn? *Educational Psychology Review*, 16, pp.235-266, 2004

- [9] A. C. K. Cheung, R. E. Slavin: The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis, *Educational Research Review*, 9, pp.88–113, 2013
- [10] L. Vygotsky: *Mind in society: The development of higher psychological processes*, Harvard University Press, 1978
- [11] H. Garfinkel: *Ethnomethodology's program: Working out Durkheim's aphorism*, Lanham, MD, Rowman & Littlefield, 2002
- [12] T. Koschmann, A. Zemel: Optical pulsars and black arrows: Discoveries as occasioned productions, *Journal of the Learning Sciences*, 18, pp.200–246, 2009
- [13] M. Kapur: Productive failure, *Cognition and Instruction*, 26, pp.379–424, 2008
- [14] B. J. Reiser: Scaffolding complex learning: The mechanism of structuring and problematizing student work, *Journal of the Learning Sciences*, 13, pp.273–304, 2004
- [15] M. Kapur: Productive failure in mathematical problem solving, *Instructional Science*, 38, pp.523–550, 2010
- [16] S. Knight, A. F. Wise, B. Chen: Time for change: Why learning analytics needs temporal analysis, *Journal of Learning Analytics*, 4-3, pp.7-17, 2017
- [17] A. Andrade, J. A. Danish, A. V. Maltese: A measurement model of gestures in an embodied learning environment: Accounting for temporal dependencies, *Journal of Learning Analytics*, 4-3, pp.18-45, 2017
- [18] E. Haataja, J. Malmberg, S. Järvelä: Monitoring in collaborative learning: Co-occurrence of observed behavior and physiological synchrony explored, *Computers in Human Behavior*, 87, pp.337–347, 2018
- [19] B. Chen, S. Knight, A. F. Wise: Critical issues in designing and implementing temporal analytics, *Journal of Learning Analytics*, 5-1, pp.1-9, 2018
- [20] M. M. Chiu: Statistically modelling effects of dynamic processes on outcomes: An example of discourse sequences and group solutions, *Journal of Learning Analytics*, 5-1, pp.75-91, 2018
- [21] N. Mercer: The seeds of time: Why classroom dialogue needs a temporal analysis, *Journal of the Learning Sciences*, 17, pp.33-59, 2008
- [22] P. Reimann: Time is precious: Variable-and-event-centered approaches to process analysis in CSCL research, *International Journal of Computer-Supported Collaborative Learning*, 4, pp.239-257, 2009