

# 教育ビッグデータをベースとした新型 e-learning の展開

## - 人間の履歴データの特徴と排除できない大きな誤差 -

寺澤 孝文\*1

Email: terasawa@okayama-u.ac.jp

\*1: 岡山大学大学院教育学研究科

◎Key Words 教育ビッグデータ, 学習意欲, データアナリティクス, アダプティブラーニング

### 1. はじめに

近年, ICT の進歩により, 大量のデータを集約・解析できる環境が整い, 機械学習 (=AI) の進歩により様々なデータの処理が人間の処理レベルに近づいている。それに伴い, 収集されるデータの解析を担う, 計算科学系の研究者を集めたデータサイエンスの研究拠点が全国的に生まれ始めている。一方で, パブリックデータの集約はなかなか進んでいない。特に, AI の開発では人間の経験(行動)データが必須であるにも拘らず, それを大規模に収集している場は残念ながら十分あるとは言えない。人の行動データを収集することはそれ自体が難しいことである。

本発表では, ICT の進歩により集約することが容易になってきている, 大量の人間の行動履歴データの特徴と影響力の強い要因について紹介し, その影響を排除する方法(マイクロステップ・スケジューリング法)について説明を加える。さらに, その技術を利用し, 手に入り始めた, 非常に精度の高い行動ビッグデータを活用し, 社会実装が始まった新型 e-learning の展開について紹介する。

### 2. 人間の行動に影響する時間軸上の条件

#### 2.1 Aさんが明日ビールを買う確率はわからない

教育情報系の企業は, かなり以前より大量の教育ビッグデータを収集している。それにもかかわらず, それらの企業から, 個人が必要とする画期的な情報サービスは生まれていない。その原因は, 人間の行動履歴データに必ず含まれる, 時間軸上の条件の影響が大きな誤差を生み出すため, ビッグデータの精度が低くなってしまったことによる。つまり, TSUTAYA の T カードにより膨大な購買行動データが収集されたとしても, そのビッグデータからは有意義な知見は出て来ない, 本質的な問題がある。

わかりやすい例では, ビッグデータの解析により A さんのビール消費量が飛び抜けて多いことが分かったとしても, A さんが明日ビールを買う確率を推定することはできない。なぜなら, A さんが今日ビールを 1 ケース購入していた場合, 一般的明日買うことはないと予想できる。人間の行動は, 「いつ」その行動を起こすのかというタイミングと評価(テスト)するまでのインターバルによって大きく変化する特徴を持ち, さらにその効果は想像を超えるほど大きく, 累積的にその後の行動に影響を与える。

鉄は 1000 度に熱しても元に戻れば同じ特性を示すが,

人間は火傷で一度熱い思いをしたらしばらく同じ行動はとらない。

そのような行動の時間軸上の条件はほぼ無限に想定でき, かつ, 人は日常の生活の中で自由なタイミングで行動を起こしている。それ故, 全人類を対象に全ての行動データが集約されたとしても, 一つとして同じ経験はなく, 一般的な行動予測を導き出そうとすれば, 時間条件の影響が大きな誤差となるため, 誤差以上に効果の大きな, いわばありきたりの予測しか手に入らない。

どんなに大量のデータが収集でき, 新たな時系列データの統計解析や AI による分析技術が進歩したとしても, 時間次元の要因に起因する誤差は事後的には排除できない特徴を持つ。

#### 2.2 時間軸上の条件に関する研究

「いつ」という時間軸上に想定される条件の影響については, 社会科学の領域の中でも, 実験心理学の領域で限定的であるが検討がなされてきた。例えば, 強化スケジュール, 分散・集中学習といった, 時間軸上におけるイベントの生起タイミングについての表現法があり, その影響について研究がなされている (e. g., Dempster, 1996 ; 北尾, 2002 ; 水野 2002)。また, 記憶研究では, 同じ学習項目が反復提示される場合の間に入れられる間隔が記憶成績に与える効果 (spacing effect: e. g., Hintzman & Block, 1973) 等はかなり研究がなされてきた。

しかし, そこで行われるタイミングの検討は, やはり限定的といわざるを得ない。例えば, 分散・集中学習の研究で複数のタイミング条件が設けられたとしても, 最終的な議論は, 分散学習条件か集中学習条件のいずれかが効果が高いという, 2 分法的な表現に帰着させられる形でなされており, 具体的にどのようなタイミングが学習効果を高めるかという議論はなされていない。繰り返しの回数や覚え方といった条件は限定的であり, 研究者が考慮することができるが, タイミングはそもそも無限に想定でき, 個々のタイミングを厳密に比較検討することは困難である。

#### 2.3 比較が難しいタイミング条件の影響

日常の学習場面で, 学習者が自由なペースで勉強する状況下でなされる反応のタイミングは, まさに無限に想定でき, 一つとして同じタイミングによる学習はなされない。また, バラバラなタイミングで学習をし

た後、特定の日にテストがなされた場合には、学習してからテストまでのインターバルもまちまちになってしまう。一夜漬けの学習効果が残っている場合と、1ヵ月ほどのインターバルでテストがなされる場合とでは、成績は全く異なる。加えて、ひとくちに学習の効果といっても、タイミングが異なる学習スケジュールの違いの何が学習効果に影響しているのか言及することすら難しい。つまり、単純に学習データを集めているだけでは、学習の効果自体を厳密に評価することは不可能といえる。それではどうすればよいのであろうか。

### 3. 人間の行動に影響する時間軸上の条件

#### 3.1 スケジュールを生成する(制御する)技術

1つの答えは、事後的に制御が困難な時間次元の条件を事前に制御することである。つまり、事前に学習やテストのタイミングやインターバルを制御した上で、反応データを回収する方法である。

学習や広告の効果については、提示したコンテンツに対する反応を回収することが可能であり、提示スケジュールは制御可能である。ただし、学習スケジュールに関しては、コンテンツは多数あり、学習者が一日当たりで費やせる学習時間は短い。この状況下で学習のタイミングとインターバルを制御するためには、様々な問題を解決しなければならない。

例えば、従来の実験計画法により、1000語の英単語の全てを4回繰り返し学習してもらい、それから1ヵ月のインターバル後のテストで学習効果を測定する実験を想定してみる。その場合、学習だけで数時間を要し、実質的に実験はできない。当然、学習を何日かに分ける必要が出てくるが、その場合、今度はテストまでのインターバルが学習内容ごとによって変わってしまい、インターバルを統制できなくなる。この例は典型的な問題であるが、多数のコンテンツを対象に連続して学習やテストを実施するための実験計画法はこれまでなかった。

寺澤・吉田・太田(2007)は、それらの問題を解決する実験計画法(スケジュールリング原理)を考案し、さらに、その原理をコンピュータシステムに実装し、実際にスケジュールに従って膨大なイベントを生起させ、それにヒモ付けて反応を記録するコンピュータシステムを開発した。さらに、一般の学校で実際に個人データを収集し、解析結果をフィードバックする一連のシステムを構築した(スケジュールリングシステムの実装の流れは、寺澤(2006:特許)などで公開されている)。

#### 3.2 学習者ごとに成績の上昇が可視化される

上記の技術により、学習者ごと、コンテンツごと(英単語ごと)に詳細なスケジュールを年単位に生成することが可能になり、そのスケジュールに従って学習・テストイベントを生起させ、それぞれに対応する反応データ(判定、反応時間等)を全て回収することを実現し、時系列条件がそろった大量の学習・テストデータを一元的に集約することが可能になった。本研究では、このデータを**高精度教育ビッグデータ**と呼ぶ。そのデータから、学習者一人ひと

りについての学習成績の上昇を解析すると、わずかな学習効果の積み重ねを個別に可視化することが可能であることが明らかになった。図1は、3名の児童が、3週間漢字学習を行った場合の、各児童の成績の変化を描き出している。児童Aの子どもは従来のテストでは、成績の上昇は得られず、相対評価では常に低い評価がなされていたが、図のように着実に成績が上昇することが明らかになった。このようなデータを個別にフィードバックするシステムも構築され、一人ひとりに、頻繁に成績の上昇をフィードバックする支援が実現された。

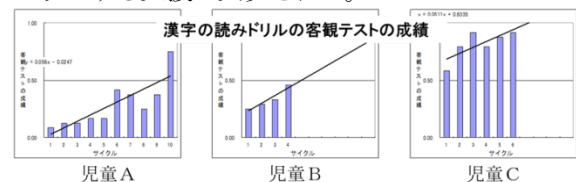


図1 3人の児童の3週間の学習成績の変化

#### 3.3 学力低位の子どもの意欲が確実に向上

図1のような個人レベルの成績の上昇を可視化している研究は世界的にもない。当初、我々も個人レベルでこのような結果が描き出されることは想定していなかった。逆に、このような想定を超えるデータが得られた事実は、時間軸上の条件の影響が非常に大きいことを示唆している。

その後、上記データをフィードバックすることが、学力低位(意欲を失っている)の子どもの主体的な意欲を劇的に向上させられることが、一般の学校で行われている支援で明確になってきた(2018年日本心理学会で発表予定)。

#### 4. 広がる社会実装

高精度教育ビッグデータをバックボーンとする新型e-learningにより、基礎学力はもちろん、子どもの主体的意欲を科学的に向上させられることが明確になってきたことで、そのe-learningの社会実装が広がっている。徳島県、岡山県赤磐市、長野県高森町、福岡県の私立高校、また、国立大学4校、私立大学4校などで年単位で学習支援がなされ、膨大な教育ビッグデータが集約され、研究者が共有するしくみも構築されつつある。

#### 引用文献

- (1) Dempster, F.N. : Distributing and managing the conditions of encoding and practice. In E. L. Bjork and R. A. Bjork(Eds.). Handbook of perception and cognition (2nd ed.). Academic Press. pp.317-344.(1996).
- (2) 北尾倫彦：“記憶の分散効果に関する研究の展望”，心理学評論, 45, 164-179(2002).
- (3) 水野りか：“分散効果の知見に基づく効果的、効率的でやる気の出る反復学習方式の考案と検証”，教育心理学研究, 50, 175-184(2002).
- (4) Hintzman, D.L. & Block, R.A. : “Memory for the spacing of repetitions” ,Journal of Experimental Psychology, 99, 70-74(1973).
- (5) 寺澤孝文, 吉田哲也, 太田信夫：“マイクロステップ計測法による英単語学習の個人差の測定”，風間書房(2007).
- (6) 寺澤孝文：“スケジュールデータ配信評価方法”，特許第3764456号