

教育支援システムにおける評価の再考：何のための評価か？

On Evaluation for Research in Educational Systems:

What is an evaluation for?

松田 昇*, 林敏浩**

Noboru Matsuda* and Toshihiro Hayashi**

*ピッツバーグ大学, **佐賀大学

*University of Pittsburgh, **Saga University

あらまし 現在も含めこれまで、教育支援システムについて多くの開発・研究が行われ、それに関わる者にとって「評価」は常に重要なファクタとして認識されている。本稿では、教育支援システムの研究における「評価」のあり方について考察する。特に、教育支援システムの開発という領域における科学的な研究の構造の中で評価の位置づけおよび役割を考察する。

キーワード 教育支援システム, 評価, research question, テーゼ, 仮説の検証

1 はじめに

本稿では、教育支援システムの研究における「評価」のあり方について考察する。教育システムの開発に従事する研究者であれば「論文のために評価は必要か?」「何をいかに評価すれば良いのか?」という疑問に誰でも悩まされるであろう。ここでは、あえて、教育支援システムを“開発”と“研究”を分けて捉える。その上で、教育支援システム開発の領域において科学的な研究構造の中で評価の位置づけと役割を考察する。

上記の視点からは、研究を遂行するために評価は必要不可欠であることが極めて自然に浮かび上がってくる。そして、何のために何を評価するのか?という問いに対する答えは、自ずと見えてくると考える。

2 教育支援システム研究における評価とは?

一般に、科学的な研究とは、research question (テーゼ)に答える営みである。そこに何らかの未知なる現象が存在し、それを科学的に解明するのが研究の醍醐味である。では、教育支援システムの研究とは何であろうか?ここでは「教育支援システムとは、教育という営みに含まれる生徒および教師の活動を支援するシステム」と広義に定義する。さらに、一般に、支援の良し悪しは、学習効果で測定できるとする。すなわち、教育支援システムの研究とは「高い学習効果を提供するシステムの開発に関わるテーゼに答える営み」であると言える。それは、すなわち「何故?」「何が?」という形式の問いに答えることである。

教育支援システムの開発など、実験的な研究では、現象そのものを見定めようとする探求的な段階から、現象を支配する理論を構築する段階へと発展するのが一般的である。教育支援システムの開発研究では、それ

は、具体的なドメインを想定したシステム作りの段階から、システムの有効性を掘り下げて解明する段階へと発展していく過程に等しい。MYCIN から EMYCIN への発展を伴う GUIDON の研究はその典型である [1]。探求的な段階では、例えば「 という特徴を持つドメインで、 な機能を持つ ITS を開発すると学習効果が高まるのは何故か?」という research question が掲げられる。この段階では、research question 自体は比較的単純であるが、その答えが記述的になり「一言で言うのは難しい」という場合が多い。一方、研究の積み重ねで、現象をより因果的に捉えるようになると、「 というスキルの定着には、 な支援形態は効果があるか?」という research question が多くなる。このような research question は、一般に問いそのものが記述的で複雑になるが、その答えは Yes および No という単純なパターンになる点が特徴的である [2]。

上述の research question には必ず答えが存在する。そして、科学的な研究では、その答えが“科学的”に裏づけられることが流儀である。多くの場合、新しい研究を提案する際には research question に対するその研究者なりの答えが想定される。その答えは、一般的には仮説 (hypothesis もしくは supposition) と呼ばれる。仮説は検証されるべきであり、すなわち、テーゼに対する答えは裏づけられるべきである [3]。

教育システムの研究では、一般的に、この仮説に対する裏づけが「評価」という形式で現れる。

3 評価はしなければならないのか?

教育支援システム開発を科学研究として捉えると、research question に対する答えの裏づけという意味で「評価」は必要不可欠である。では、その「評価」と

は何か？システムの挙動を評価するのか？生徒の学習を評価するのか？システムの開発コストを評価するのか？これらの問いは研究に依存する．すなわち，評価は research question に対する仮説の裏づけであるから，「仮説の裏づけ」に必要な評価を行うべきである．「システムの評価」＝「学習実験」という短絡的な偏見は，誤謬であることに気づく．もっとも，科学研究での仮説の捉え方と検証方法は，長い間，哲学的に議論され続けている [4] ので，本稿ではこれ以上立ち入らない．

我々は，ここで哲学的な議論をするのではなく，身近な例から，検証としての評価のあり方を学ぶべきであろう．例えば，ITS を実際の学校現場に導入し学習効果が測定された大規模な実験 [5, 6] などがある．

さて，教育支援システムの研究において，評価は検証のみのツールではない．例えば，質的研究で狙いとされる「現象を掘り下げることにより，原理を見出す」というスタイルの探求的な研究では，評価は，例えば，「
という教育場で特徴的な現象は何か？」という open end な research question に答えるために用いられることがある．例えば，Schofield は，学校現場に ITS を導入した際の教師および生徒の触れ合い方の変化を観察した [7]．Shute は ITS による学習評価を通して，General Intelligence と学習効果の関連を調査した [8]．このような研究では，評価そのものが研究の対象またはツールなので，「教育支援システムの研究に評価は必要か？」の問いと並列して議論するのは，少々場違いかもしれない．しかし，目的に応じて評価の仕方が変わることを理解する上では，良い例であろう．

4 教育システムの評価は難しいのか？

答えは Yes でもあり No でもある．Research question に対する仮説を検証する立場での評価は，合目的的であるという意味で，その捉え方は簡単であると言える．しかし，具体的に何をどのように評価するのか？という話になると，実験計画に関わるデリケートな議論であり，一朝一夕で片付くとは思えない．

探求的な評価は，いわば，カオスから“意味のある”何らかの法則を見つけ出す試みであり，本質的な困難が伴うと考える．「教育支援システムの評価は難しい」と一言で片付けられる場合，多くは明確な research question を持たずにシステムを構築して，「さて，論文を書くために何を評価しようか？」と悩んでいるように見受けられる．少々データは古いが，例えば，2000 年に教育システム情報学会に掲載された原著論文において，検証的な評価が行われている論文は 2 割に過ぎない (18 本中の 4 本)．3 割の論文は，research question を見出すことすら著しく困難である．残りの論文は，評価と

research question が対応しているとは考えにくい．

5 おわりに

教育支援システムの開発における評価は，その目的により捉え方が異なる．本稿では，それらを大きく「検証のための評価」と「探求のための評価」に分けて考察した．科学的研究を遂行する上での research question に対する仮説 (supposition) は検証されるべきであり，それは，一般に“評価”として実装される．一方，探求的な研究では，学習活動が評価の対象になる．すなわち，学習活動の評価を通して，未知なる現象を解明すること研究の目的である．そのような探求的な研究を通して得られた知見は，新たな教育支援システムの開発に示唆を与えるであろう．

最後に，この論文のタイトルに対する答えにお気づきであろうか？筆者らは，それを「学者の好奇心」であろうと思いたい．

参考文献

- [1] Clancey, W. J.: “*Knowledge-Based Tutoring: The GUIDON Program*”, Cambridge, MA: MIT Press(1987).
- [2] Cohen, P. R.: “*Empirical Methods for Artificial Intelligence*”, Cambridge, MA: MIT Press(1995).
- [3] Van Wageningen, R. K.: “*Writing a Thesis: Substance and Style*”, Englewood Cliffs, NJ: Prentice-Hall(1990).
- [4] Tweney, R. D., Doherty, M. E. and Mynatt, C. R. (Eds.): “*On Scientific Thinking*”, New York, NY: Columbia University Press(1981).
- [5] Anderson, J. R., Corbett, A. T., Koedinger, K. R. and Pelletier, R.: “Cognitive tutors: Lessons learned”, *Journal of the Learning Sciences*, Vol.4, No.2, pp.167-207(1995).
- [6] Koedinger, K. R., Anderson, J. R., Hadley, W. H. and Mark, M. A.: “Intelligent Tutoring Goes to School in the Big City”, *Journal of Artificial Intelligence in Education*, Vol.8, pp.30-43(1997).
- [7] Schofield, J. W., Evans-Rhodes, D. and Huber, B. R.: “Artificial intelligence in the classroom: The impact of a computer-based tutor on teachers and students”, *Social Science Computer Review*, Vol.8, No.1, pp.24-41(1990).
- [8] Shute, V. J. and Glaser, R.: “A Large-Scale Evaluation of an Intelligent Discovery World: Smithtown”, *Interactive Learning Environments*, Vol.1, No.1, pp.51-77(1990).