

eラーニングにおける多様なデータを活用した質保証と 評価のフレームワーク

松田 岳士・合田 美子・玉木 欽也

本稿では、青山学院大学総合研究所eラーニング人材育成研究センター（eLPCO）が取り組んでいるeラーニングの質保証の枠組みと、その特徴を先行研究と対比させながら論考した。先行研究を調査した結果、第一に、eラーニングに関する既存の評価や質保証は、おおむねチェックリスト方式を取っており、個々のeラーニングプログラムの改善を目的として絶対評価を実施するものと、複数のeラーニングプログラムの比較を目的として相対評価を実施するものがあることが判明した。また、第二に、先行研究の中には、複数の品質保証レベルを設定して、全く異なるアプローチを組み合わせたフレームワークもあった。

これらの先行研究の利点や課題を勘案して、eLPCOでは、質保証の枠組みにチェックリスト方式とデータマイニング方式を併用することとし、4つの評価フレームを採用して質保証の活動を行なっている。それらは、各授業の評価（フレーム1）、データマイニングを用いた新たな評価モデル研究（フレーム2）、eラーニング専門家認定プログラムとしての評価（フレーム3）、評価活動自体の評価（フレーム4）である。本稿では、そのうちフレーム1の取組みと総体的な評価として実施されるフレーム2の取組みの現状について報告した。

キーワード

品質保証の枠組み、チェックリスト方式、データマイニング方式、個別授業の評価、総体的な評価

1. はじめに

eラーニング、とりわけ非同期分散環境で実施されるeラーニングは、これまでの対面授業や通信教育と比較すると、学習者にとっても教育を提供する側にとっても様々なメリットを持っている。松田（2006）は従来の通信教育と比べたeラーニングの特長を、次の4点に整理している。

- ① 双方向コミュニケーションのコストとスピード改善
- ② マルチメディアを用いたコンテンツ
- ③ 豊かな学習資源が眠っているインターネット
- ④ 多くの学習行動や学習効果の自動保存

これらの特長をeラーニングの質保証という観点から考察すると、④が重要であることがわかる。これまでの通信教育では、学習者が何をどのように学んでいるかについて、ほとんど知ることができなかったばかりか、学習支援のために教育を提供する者がタイミング良くコミュニケーションを取ることも難しかった。しかし、eラーニングでは、従来の通信教育はもちろん、大規模な

対面授業と比べても、学習者の進捗状況、学習スタイル、学習効果をより詳細に知ることができる。

特に学習管理システムの発達とコンテンツの標準化によって、現在のeラーニングシステムは、教材へのアクセス状況、テストの成績、学習者とメンタやチュータとのコミュニケーションなど様々な学習履歴を保存、管理するデータベースとして機能している。このような状況が実現されている以上、学生の満足度をアンケートで質問し、その平均値が一定レベルを超えていれば、eラーニングが成功であり、教育・学習の質保証がなされるとみなしてしまうのは、少なくともeラーニングの質保証としては、不十分である。

それでは、eラーニングにおける質保証はどうあるべきなのであろうか。eラーニングプログラムでは、学習プログラムを評価できる指標の幅が広がっただけなのであろうか。このような疑問に答えるため、青山学院大学総合研究所eラーニング人材育成研究センター（Research Institute for e-Learning Professional Competency：以下eLPCO）では、2005年4月の設立当初より、eラーニングにおける質保証への取組みを担当する研究部会としてTask Force (TF) 13を設置し、eラーニングプログラム評価研究と評価者育成カリキュラムの開発をテーマに研究を続けてきた（松田 2006）。

本稿では、TF13の実際の活動に沿ってeLPCOへの質保証の取組みを紹介する。具体的には、eラーニングの質保証に関する先行研究を概観し、これまでのTF13の評価基準や評価フレーム開発活動について報告するとともに、2006年度前期に実施されたeラーニング授業の評価活動のあらましを示す。

2. eラーニングの質保証の特徴

eラーニングの質保証と評価について研究するために、TF13がまず行なったのは、そもそもeLPCOにおけるeラーニングの質保証とは何であるかを検討することであった。

そこでTF13は、評価活動に取り組むのに先だって、国内外の評価機関、団体、研究者による既存の評価基準をレビューした。そのうち本稿でとりあげるのは、アメリカ、ヨーロッパ、日本における基準である。

最初に、アメリカのアクレディテーション団体の評価項目に加えて、同じくアメリカの独立系シンクタンクによって作成されたQuality On the Line (The Institute for Higher-Education Policy 2000)、Alfred P. Sloan財団のスポンサーシップによって遠隔教育実施機関の評価のため設立されたコンソーシアム、Sloan-Consortium (Sloan-C)によるElements of Quality Online Education (Bourne and Moore 2003)、そして遠隔学習としてのeラーニングに焦点を当てたPorter (1999)の先駆的な研究、多角的な評価手法を模索しているHorton (2001)を検討する。

次に、ヨーロッパを中心に設定されている評価基準をとりあげる。ここでは、CEN/ISSS (2003)が作成したCWA Quality Assurance Standards、およびISO/IECのJTC1 SC36で協議が続いているReference Framework for the Description of Quality Approaches (RFDQ)の2003年度報告書における評価基準を概観し、続いて、日本を中心としたeラーニングの評価・認証基準として、Asia e-Learning Network (AEN 2005)のコンテンツ、サービスの品質基準について考察する。これら米欧日の評価基準は様々な視点から教育の質を保証しようと試みているが、それらには共通する課題があるので、最後にその課題を整理する。

まず、アメリカ地区アクレディテーション団体の共同ガイドラインBest Practices for Electronically Offered Degree and Certificate Programs (North Central Association Commission on Institutions of Higher Education 2000)をみていく。このガイドラインは、5部構成であり、(1)教育機関の置かれている文脈と貢献 (Instructional Context and Commitment)、(2)カリキュラムと教授法 (Curriculum and Instruction)、(3)教員への支援 (Faculty Support)、(4)学生への支援 (Student Support)、(5)プログラム評価と学生の成績評価

(Evaluation and Assessment)に分かれている。

それぞれの評価項目は、さらに4から10の下位項目に分かれており、さらに一つ一つの下位項目がいくつかの具体的なチェック項目として質問の形式をとっている。例えば、プログラム評価の自己評価に関する下位項目は、次のように記述されている。

表1 Best Practices for Electronically Offered Degree and Certificate Programs (部分)

| |
|---|
| <p>5e. 教育機関は、以下のような目的のために継続的な自己評価プログラムを実施する。教育プログラムの改善、教授法改善のためのより効果的な技術使用、意図した学習成果を達成するための進捗、ドロップアウト率の低下、リソースの効果的使用、内外の支援者に対する教育機関の明確なサービス向上。自己評価プログラムとその結果は、現在行われている自己評価プロセスに反映され、教育機関とアカデミックプログラムに責任を持つ者の将来的な計画を告知するために使用される。</p> <ul style="list-style-type: none"> ・教育機関が現在行っている評価と改善のプログラムはどのように開発され実施されているか。 ・そのプログラムは、学習成果改善、学生の維持、リソースの使用、中心的な支援者へのサービスの本質的なカテゴリをカバーしているか。 ・そのプログラムは、学術的な資格を有する者を適切に含んでいるか。 ・既存の教育プログラムやコースの見直しと修正のために教育機関はどのような機構を持っているか。 ・教育プログラムの評価は教育機関の計画立案にどのように影響するか。 ・どの支援者が改善のための計画立案に積極的に関与することになっているか。 ・評価プロセスはこれまで測定可能な成果をあげてきたか。 |
|---|

(North Central Association Commission on Institutions of Higher Education 2000)

このような構造を持つガイドラインの特徴は、教員、学生双方に対する支援体制を細かく具体的にチェックしていることである。先述した主要5項目は、直接教員や学生への支援に関するセクションばかりでなく、カリキュラムと教授方法や評価のセクションにも及んでいる。

例えば、カリキュラムセクションでは「主たる教員以外にメンタなどを用いる場合、学術的な資格を有しているか」、評価セクションでは「教育プログラムによる経験についての学生の全般的な満足度調査；その調査には学生から見たコストの得失を反映させる」といったチェック項目が見うけられる。

次にQuality On the Line (QOL)は、高等教育機関が提供するオンライン教育のベンチマーク（比較のための基準）を開発する目的で作成された。その評価基準は次のような7つのカテゴリと24の基準に分けられている。

表2 Quality On the Line の評価基準

| |
|--|
| ①教育機関からの支援基準 (Institutional Support Benchmarks) |
| ・文書化された技術プラン (セキュリティ含む) |
| ・配信システムの信頼性 |
| ・集中管理システムがインフラへのサポート提供 |
| ②コース開発基準 (Course Development Benchmarks) |
| ・コース開発のための最低限の基準についてのガイドライン |
| ・指導用の教材を定期的に再検討 |
| ・学習者自身に分析, 総合, 評価に関わるよう求めるコース |
| ③教育/学習過程基準 (Teaching/Learning Process Benchmarks) |
| ・学習者が教員や他の学習者と交流するための様々な方法 |
| ・学習者へのフィードバック |
| ・調査のための適切な方法を指導 |
| ④コース構造基準 (Course Structure Benchmarks) |
| ・学習者の動機, 技術使用などに関する事前アドバイス |
| ・補助コースの情報提供 |
| ・図書館情報へのアクセス |
| ・課題提出と教員の評価についての合意 |
| ⑤学習者支援基準 (Student Support Benchmarks) |
| ・プログラムに関する必要な情報提供 |
| ・資料入手のための訓練や情報 |
| ・コース期間中の技術的援助 |
| ・質問への適切, 迅速な応答 |
| ⑥教員支援基準 (Faculty Support Benchmarks) |
| ・コース開発における教員への技術支援 |
| ・オンラインコースへの移行期間中の補助, 給与支払い |
| ・コース期間中のインストラクタからの訓練と補助 |
| ・学生の電子データ使用に関する問題への文書による資料提供 |
| ⑦プログラム評価/成績評価基準 (Evaluation/Assessment Benchmarks) |
| ・複数の方法を用いた, 特定の基準を適用した評価プロセス |
| ・学生数, コスト, 革新的な技術についてのデータを使用 |
| ・学習効果目標の定期的な見直し |

(The Institute for Higher-Education Policy 2000)

QOLの成立過程を追っていくと, 作成の過程で, コンテンツの質より教員や学習者のコミュニケーションに関する評価項目の重要性の方が高いと判断されたことがわかる。その結果, 上記のように教授・学習過程の支援に多くの項目が充てられることになった。

これに対して, Sloan-Cの品質保証は「だれでも, いつでも, どこでも多様な教育にアクセスできる」(Mayadas et. al 2003) ことを目標としている。そのため, Sloan-Cでは非同期学習ネットワークによるコミュニケーションの品質を重視する立場を取る。

Sloan-CのElements of Quality Online Educationでは,

具体的には評価に関する5つの要素を提示している。それは, (1)学習効果 (Learning Effectiveness), (2)費用効果 (Cost Effectiveness), (3)アクセス (Access), (4)教員の満足度 (Faculty Satisfaction), (5)学習者の満足度 (Student Satisfaction) である。

これらの5要素のうち, 学習効果に関する項目は, 十分な設備, 施設, 制度が整っているかという質問が並び, アクセスに関する項目では, インタラクティブ性をどこまで利用できたかをチェックしている。さらに, 教員や学習者の満足度では, 教育機関から教員や学習者への支援が確認されている。

また, Sloan-Cでは5要素ごとに部会を設けて事例を収集し, 毎年評価項目を修正している。さらに, 要素ごとに優れた実践を表彰する事業も行っている (Sloan-C 2004)。

表3 遠隔学習コース評価リスト (部分)

| 評価項目分類 | 具体的チェック項目 | |
|---------|--|---|
| 一般的評価項目 | コース内容 | ・コースの期間はどのくらい続くか ・教育・訓練者は誰か ・他の人々がこのコースをこれまでどのように評価してきているか |
| | 学位プログラム | ・遠隔学習を介していくつのコースを履修可能か ・キャンパスで提供される他の学位プログラムにどれくらい匹敵するか ・教材はどの程度頻繁に更新されるか |
| | 時間枠 | ・活動を完了するための時間制限はどれくらいか ・どれくらい速くコースを完了できるか ・どれくらいの頻度でコースが提供されるか |
| | 費用 | ・プログラム全体の費用はいくらか ・教材のための費用はいくらか ・(割引など) 料金体系はどのように決められているか |
| 技術的要求 | ・コースはどこで受講できるか ・どのような種類の遠隔学習環境が用いられるか | |

Porter (1999)

次にPorter (1999) は, 日本ではまだインターネットを用いたeラーニングが一般的でなかった1997年に, インターネットによる遠隔学習に関する総合的な解説書を記し, その中で学習者, 授業提供者両者にとっての適切なコースの基準を提示した。Porterはその中で, 資金調達の方法からコースの設計時に注意すべき問題までを詳

細に論じている。

Porterは評価項目の設定に際して実証的なデータを示しているわけではなく、一つ一つの項目にどの程度妥当性や信頼性があるのか明らかでないが、eラーニングに関する問題点を体系的に整理したという意味では、先駆的、かつ資料的価値がみとめられるうえ、注目した問題のいくつかはTF13における研究とも関係がある。なかでも成人学習者のためのコース評価チェックリストは、網羅的であるばかりでなく、学習者から見たコースの適切さとは何かを考察する際の参考になった。表3にその一部を示した。本稿3.1で述べるフレーム1の評価項目は、Porterの評価リストをカバーするように設定されている。

一方、Horton (2001) はeラーニングプログラムの評価に特化した文献で、Kirkpatrick (1998) の4段階評価を応用した評価スキームを手順化している。Kirkpatrickの4段階評価は、トレーニングコースの評価基準として広く用いられているもので、第1段階である受講直後の学習者の「反応」、すなわち満足度から第2段階の「学習」、つまり学んだ知識やスキルの程度、第3段階の「パフォーマンス」、学んだことが仕事や行動にどのように反映されたか、第4段階の「結果」、つまり結果的に組織にとっての利益はどの程度であるかを順に確認していくものである。

Hortonの評価の特徴は、従来実施困難であったKirkpatrickの第3段階、第4段階の評価を実現するための具体的な方策をいくつか提示していることと、eラーニング授業提供者が独自性のある評価報告書を書くことを目標として、評価者自身が各項目に重み付けすることによって質的評価を点数化できる基準を示していることである(表4)。

実際の評価における判断レベルは、学習プログラムごとに違っているはずであるというHortonの考えは、TF13においても個々の授業の評価基準設定の元になった。それは、次のような考察に基づいている。

まず、教育プログラム、特に高等教育における授業は、大量生産による製品作りとは本質的に異なっている。しかも、非同期分散型のeラーニングのように、いつでもどこでも学べる授業は、教育を提供する側からみると学習者がいかなる時に、いかなる場所からアクセスするか統制しにくいばかりでなく、教育提供者が学習行動を過度にコントロールすることによって却って利点が失われることにもなりかねない。eラーニングの授業環境は学習者側の環境に依存する部分が大きく、基本的にケースバイケースの運営にならざるを得ない。そのようなeラーニングプログラムを評価するのに、画一的な評価基準で対応しようとするには矛盾が生じるのである。

これらの評価基準に対して、Quality Assurance StandardsおよびRFDQはeラーニングの教育過程に関する標

表4 eラーニングプロジェクトの評価基準(部分)

| 項目 | 重み | 得点* | スコア |
|--|----|-----|-----|
| ビジネス コースの目的は学習者の目的と一致しているか コース料金は学習者の経済事情に一致しているか 学習者のスケジュールに合っているか | | × | = |
| 技術 学習者が持っているコンピュータで受講できるか 新たなソフトをインストールしなくてもよいか 既存のLMSで受講できるか | | × | = |
| コンテンツ 教材は正確で最新のものか 学習目的に合った範囲をカバーしているか | | × | = |
| インストラクショナルデザイン コースのタイプは学習目的に最適か 学習者の理解を助ける提示順になっているか | | × | = |
| ユーザビリティ オンラインの補助のみでコースを開始できるか 問題や疑問がある時に何をすればよいかわかるか | | × | = |
| メディア 画像は適切に使用されているか テキストは学習者が完全に理解できるレベルか | | × | = |
| モチベーション 学習者が何を学べるかが最初に明示されているか 学習者にとって適切な難度か | | × | = |

* 評価結果を-5から+5、0から100のようなスケールで表わしたものである

(Horton 2001)

準規格作成の観点から形成されている。CENもISOも国際的な標準化団体であり、両者とも教育プロセスを標準化することによって客観的な品質保証を達成しようと試みている。

表5はISSSの評価基準によるeラーニングコース評価の一例であるが、教育プロセスを記述するスキームを用いて、教育プロセスの比較ができるようになることを目的とするのが標準化の発想であることがわかる。なお、表5で×印が打たれている項目は、「その項目に該当する」という意味であり、+、○、-は表中にあるように、支援の程度を3段階で表したものである。

このような標準規格を用いた評価は、QOLのような評価基準とは根本的に出発点が異なっている。QOLで

表5 ドイツEssen大学の学習モデル評価

| | | |
|-----------------------------------|---|--|
| 名前 | エッセン学習モデル (ELM) | |
| 説明 | ELMは一般開発モデルであり、開発者、教育者、様々なレベルの活動を行うユーザを支援する。(以下略) | |
| ソース | http://elm.wi-inf.uni-essen.de | |
| 日付 | 2002-01-03 | |
| 方法 | | 成果指向 |
| | × | プロセス指向 |
| ターゲットグループ | + | 開発者 |
| | ○ | 学習者 |
| | + | 教員 |
| | + | コンテンツプロバイダ |
| | + | メディアデザイナー |
| プロセス +十分な支援 ○部分的支援 -支援なし | ○ | 組織的支援 (IT分析, コンテンツ分析) |
| | + | コース開発 (カリキュラムデザイン, 知識獲得, ユーザモデル形成, 学習順序デザイン) |
| | + | 教育/学習 (コミュニケーションデザイン) |
| | + | コース構造 (コースデザイン) |
| | + | 学生支援 (学習目的, コース情報) |
| | ○ | 教員支援(カリキュラムデザイン) |
| | + | プログラム評価, 学習成果評価(評価デザイン) |
| 基準 | | |
| 標準 | LOM, SCORM, ダブリン・コア | |

(CEN/ISSS 2003)

は、評価対象となるeラーニングコースが「事前に想定した一定の条件を満たしているかどうか」という任意の評価基準との一致度を確認する手順をとるのに対して、標準規格はそのコースが「どのように他のコースと比較、検討できるか」という観点を持っているので、コースの記述とコース同士の照合基準として働くといえる。

同様の発想に基づいて、チェック項目を作るために必要な表示項目を幅広くカテゴライズしたのが、AENの品質保証スキームである。AENの品質保証の枠組みでは、表6にあげたような5つのクラスが設定され、それぞれ評価の対象や評価活動が定義された。

AENの品質保証が特徴的である点は、他の評価基準で、一連の評価の枠組みにまとめられてしまいがちなeラーニングにかかわる様々な評価を、クラスという概念を用いて、評価の主体、対象となる現象、方法などの要素ごとに全く異なるコンセプトを用いて分割したことである。AENの品質保証概念は、アジア諸国のeラーニングの専門家が日本を含む各国の調査に基づいて形成したも

表6 AEN WG4の品質保証概念

| | 保証主 | 保証対象 | 評価レイヤ | 保証内容 |
|----------------|-------------|----------------|--------|--------------|
| クラス1 組織品質 | 経営者 | 全社的品質活動 | 組織体 | 組織マネジメント |
| クラス2 プロセス品質 | プロジェクトマネージャ | 開発のためのプロセス活動 | プロセス | 開発 |
| プロジェクト | | 開発のためのプロジェクト活動 | 組織体 | プロジェクトマネジメント |
| クラス3 製品品質 | 担当マネージャ | 提供機能 | 教授 | 機能 |
| クラス4 使用品質 | 担当マネージャ | 実際の教授にともなう操作 | 教授 | インタフェース |
| クラス5 学習品質 | 担当マネージャ | 学習活動・学習者の能力推移 | 学習・学習者 | 学習の効果性 |

(AEN 2005)

のであるので、アジアにおける現実的な質保証の枠組みを提案していると言える。

一口にeラーニングの質保証と言っても、学習成果の保証からビジネスの成果保証までを一種類の方法に依存して行なうことには無理がある。AENの示した5クラスは、個々のeラーニングプロジェクトを絶対評価するとしても、複数のeラーニングプロジェクトを比較するとしても、質保証の枠組みが複数必要であることを示唆している。

3. eLPCOの質保証の枠組み

ここまでみてきたような先行研究や、既存の質保証基準を勘案して、TF13は、eLPCOが実施するeラーニング授業の成果を評価するには、少なくとも次の4つのフレームが必要であるという結論に至った。すなわち、各授業の評価(フレーム1)、データマイニングを用いた新たな評価モデル研究(フレーム2)、eラーニング専門家認定プログラムとしての評価(フレーム3)、評価活動自体の評価(フレーム4)である。このうち、フレーム1、3、4の違いは評価の対象であり、フレーム1、3、4とフレーム2の違いは評価の方法である。

まず、フレーム1、3、4の違いについて確認すると、図1に示したように、評価の対象となる範囲や活動が異なる。次に、フレーム1、3、4では、対象こそ異なるものの、基本的にはチェックリストを用いるのに対して、フレーム2は、あえてチェックリストを用いない評価方式を取っている。その理由は、次の通りである。

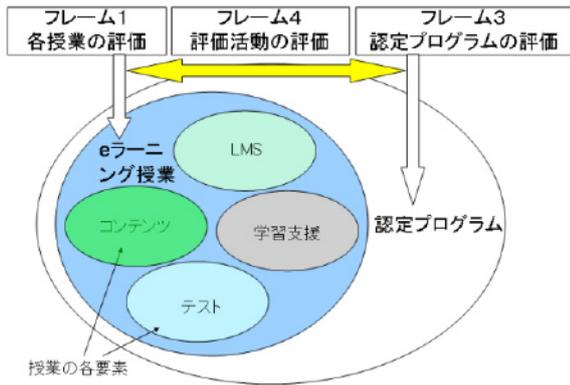


図1 評価・質保証の対象

前章までで述べたように、教育の品質保証を模索する研究は、多くが単一の評価基準に基づき、さらに、教育環境、用いられている技術、コンテンツ等を評価することを目的としている。そのため、主に教育機関の体制、コンテンツ等に関するチェックリストを用いたアプローチを取ることになる。このような方法は、eラーニングの特徴を考慮すると、少なくとも次のような課題を抱えている。

① チェックリストにない項目は評価できない。

当然のことであるが、チェックリスト方式の限界は有限の項目で形成されるチェックリストを用いることにある。しかし、事前に完全なチェックリストを作成できる手法が開発されていない限り、どのようなチェックリストにも漏れがある可能性を否定できない。Horton (2001) のリストでは、評価項目の重み付けまで可能であるものの、この限界が解消されるわけではない。

このようなチェックリストの限界を補う方法の一例として、Tergan and Schenkel (2003) は、学習者を実際の学習環境に置いて、段階的にチェックリスト作りに加わらせるアプローチ (Learner-centered Checklist Approach) を提案しているが、全てのeラーニングプログラムでこのような評価を実施することは困難である。

② 「何を学ぶか」と「どのように学ぶか」を切り離れたアプローチになる。

チェックリスト、特に標準化に関するリストは、どのようなeラーニング環境を提供するかという点に関しては有効に働き、コンテンツの機能や構造も確認できるかもしれないが、高度に専門的でしかも幅広い高等教育の学習内容については踏み込めない。

インストラクショナルデザインの評価プロセスにしても、あくまで教育提供者側を中心としたものであり、教員や教育機関が教えたいことを効率よく教えることができるかどうか、という意味における学習者とのギャップ分析に留まっている。

③ 「点」の評価になり、過程を評価するのが難しい。プログラムの終了前後のある特定の時点で評価データ

を収集してチェックリストと照合するケースが多く、コンピュータ支援による協調学習 (Computer-Supported Collaborative Learning: CSCL) のような複雑な学習活動をするコースでは、チェックリストを用いてコース途中の学習過程を評価することは難しい。

④ 学習者=消費者という認識に基づいている場合が多い。

チェックリスト形成時にマーケティングや品質管理の手法を用いる他、特に満足度の評価や製品としてのコンテンツの品質評価に関して、学習者が消費者のようなモチベーションや判断基準を持ち、消費者として行動するという前提に基づく研究がある。しかし、いくつかの研究が指摘しているように、学習者はしばしば消費者とは異なる価値観や態度を示す (Grocchia 1997, Palumbo 2003)。

また、学習者=消費者という認識の背後にあるのは教育=商品であり、教育に関しても市場原理が働くという前提であるが、教育は通常の商品と比べると公共性が高く、無条件にこのような認識を是認することには問題がある。

⑤ 評価結果を活用することが不可能なケースがある。

重要なことであるが、教員や学習者がどのようなeラーニングコースを望むかについて様々な角度から作成した質問項目を並べた理想のチェックリストを作成し、十分信頼性のある回答が得られたとしても、結局は「質の高いコンテンツが回線の帯域に関係なく使用できて、24時間学習やシステム管理をサポートしてくれる教授体制が整っている安価なコース」などという至極当然であるが実施不能な答えになってしまい、設計に役立つ情報は得られない。

次に、これらの議論をふまえて、4つの質保証・評価フレームの概要を述べる。

3.1 フレーム1：各授業の評価

このフレームによる評価は、先行研究とeLPCOの提供する授業やその環境を分析して、評価項目を分類し、評価のためのチェックリストを作る方法を用いる。このフレームによる評価の目的は、個々のコンテンツや学習プログラムの改善であるので、具体的なチェックリストを用いたアプローチが有効、かつ不可欠であると判断した。問題は、どのような分類に基づいて項目を整理するか、実際のデータ収集方法をどのように定めるかであるが、少なくとも次の5項目を評価の対象に含むことについては、TF13メンバ各自が持ち寄った素案に何らかの形で含まれており、大きな異論もなかった。

- 成績を含む直接的な学習効果
- 学習支援体制・運営体制に関する評価
- 学習活動

• 授業・コンテンツの設計

これらが大項目として、それぞれ満足度などの具体的な項目を設定していくと、ある程度のコンセンサスが得られる全授業共通のチェックリストとなるので、まずはそのようなチェックリストを作成し、授業ごとに重み付けと授業独自の評価項目を加える作業を行なった。

表7に、整理後の評価項目を記した。この項目に基づいて、実際のアンケート案が作られ、正規授業開始前に行なわれた実証実験でアンケートデータを収集し、その結果を分析してアンケートの質問項目を決定した。

表7 全授業共通評価項目

| 大項目 | 小項目 |
|----------|----------------|
| 学習効果 | 学習効果測定方法 |
| | 学習効果測定的设计プロセス |
| | テストの作成 |
| | 学習効果のレベル |
| | 学習効果のフィードバック |
| | 自己効力感 |
| | 修了率 |
| プロジェクト体制 | eラーニングシステム運営体制 |
| | 学習者支援体制 |
| | 教員支援体制 |
| | 修了率向上への取組み |
| | プロジェクト体制への満足度 |
| 学習活動 | ログイン頻度 |
| | 教材アクセス |
| | コミュニケーション |
| | 学習・教育活動への満足度 |
| 授業設計 | コンテンツ開発 |
| | 対面：eラーニング比率 |
| | 授業設計への満足度 |

3.2 フレーム2：データマイニングによる評価モデル形成

前章で述べたようなチェックリスト方式の問題点を解決する評価の方法として、eラーニングの過程で発生するデータを用いるデータマイニングの方法を開発することが可能である。この方法では、eラーニングコースの実証実験や実施によって蓄積されていく様々なデータを分析することによって評価のモデルそのものを形成していく。

表8に示したように、チェックリスト方式が、いわば事前に定めた評価指標に対する足し算・引き算による評価であるのに対して、データマイニング方式は結果のデータを因数分解する評価である。チェックリスト方式の限界が、チェックリストの項目自体にあるとすると、

表8 評価の手法と特徴

| | 相対評価 | 絶対評価 |
|----------|-----------------------------|-----------------------------|
| チェックリスト | 複数授業に共通する項目の評価シート | (特定授業の) 細分化した目標項目との合致度 |
| データマイニング | 複数授業に共通にみられるデータの検証的な比較モデル形成 | (特定授業の) 探索的データマイニングによるモデル形成 |

データマイニング方式の限界は、使用可能なデータの範囲にある。

データマイニング方式の評価では、入手可能なデータに注目することになり、今後、学習履歴等の幅が広がることによって、評価モデルが進化していく可能性がある。したがって、データマイニング方式は、評価モデル研究の新たな方向性を示すものであるとも言える。もちろん、このような研究を進めることと、フレーム1による評価を実施することは矛盾するものではない。

フレーム2の評価をさらに詳しく紹介すると、データマイニングを使用して、以下の2つの角度から研究を推進している。すなわち、①幾つかの科目をまとめて学習効果を総体的に評価する研究と、②研究部会構成員の研究興味をテーマとした個別テーマの研究である。

総体的評価活動は、学習効果、終了率、ドロップアウト等のeラーニングを利用した学習の特徴をふまえ、eラーニングと学習活動と効果の関係を明確にしようとしている。研究興味によるテーマを絞り込んだ評価では、学習者の多様性、授業設計、学習管理システム、授業形態、教材の種類、コミュニケーション方法等に注目し、細分化された側面からeラーニングの学習効果を明らかにすることを目指している。

3.3 フレーム3：認定プログラムとしての評価

これは、2008年3月に終了する現代的教育ニーズ取組み支援プログラムが成功であるかどうかを評価するというレベルである。つまり、eLPCOが提供するプログラムによって学習し、認定される5職種の専門家(インストラクショナルデザイナー、コンテンツスペシャリスト、インストラクタ、メンタ、ラーニングシステムプロデューサ)の認定そのものを検証する評価である。

例えば、所定の科目を履修して最終のテストに合格し、インストラクショナルデザイナーとして認定された学生が、実際には「知識があっても仕事はできない」状態になっていないかどうか確認することは、重要な評価の視点である。したがって、このフレームは、認定プログラムとして、結果的に育成、認定したスキルの総体を評価していると言える。

また、このような視点は、資格としての判定システム開発にとっても大切であるが、すぐに評価の結果を入手

することができない事情もあり、今後TF13に課せられた課題として引き続き検討することになった。

3.4 フレーム4：評価活動自体の評価

品質保証としての評価活動を評価するという、メタ評価の観点であり、今後評価活動自体を改善していくために必要である。チェックリスト方式を取る予定であり、現在作成している項目はAEN WG4の成果を参考にして（表9）。

表9 評価活動の評価項目（部分）

| 評価項目 | 活動レベル | 使用データ |
|------|---|---------------|
| 評価指標 | 1. 理論的な背景を持ち、組織に適応させた評価指標を用いている | 評価担当部会議事録, 資料 |
| | 2. 理論的な背景を持つ評価指標をそのまま用いている | |
| | 3. 理論的な背景を持つ評価指標を用いていない | |
| 評価機関 | 1. 外部の評価の専門家が評価を実施している | プロジェクト議事録・規則 |
| | 2. 体系的な評価スキルを持つ評価者が評価を実施している | |
| | 3. 体系的な評価スキルを持たない担当者が評価を実施している | |
| | 4. コンテンツ担当者（設計者、開発者またはインストラクタ）同士が相互評価を行っている | |
| | 5. 特に評価を行うものはない | |
| 評価時期 | 1. コンテンツの設計時に評価計画を立案している | 評価担当部会議事録, 資料 |
| | 2. 評価計画は特に立てていない | |

4. フレーム1事例：2006年度前期授業の評価

2006年度前期の授業を対象とした評価のうち、ここでは、eラーニング専門育成プログラムに参加した学生全員にとっての必修科目である「入門科目」（3科目）の評価の実際を記述する。フレーム1に則った個別の授業評価は、図2のような手続きを経て実施された。

TF13では従属変数を総体評価と満足度にとり、そのほかの入手可能なデータを独立変数に取るなどして探索的に評価モデルを作っていくアプローチが提案された。評価対象は各授業というよりプログラム全体である。

既に、チェックリストの項目決定までの流れは、3.1で述べた。それに続いて、TF13では授業の終了後、すみやかに評価報告書を執筆した。さらに、これを速報版として教員（インストラクタ）、インストラクショナルデザイナー、メンタに提示して、報告書への反応を収集し、

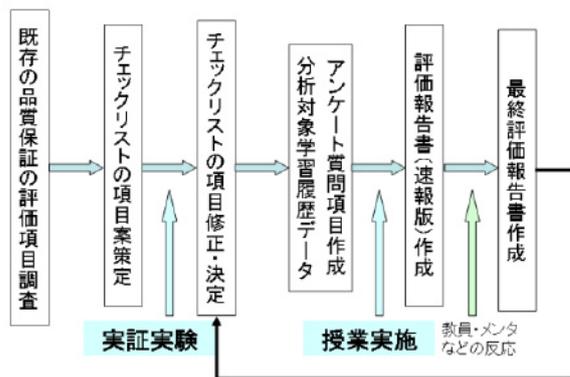
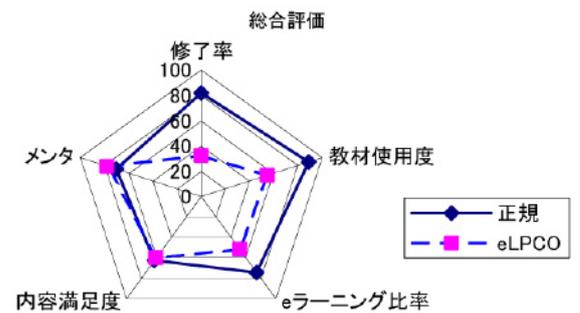


図2 各授業の評価の流れ

それを反映させて最終報告書としてまとめた。

最終報告書は、いずれも1. 評価の目的, 2. 評価対象授業の概要, 3. 主な結果, 4. 特徴的なアンケート結果, 5. TF13（評価担当部会）による総合コメント, 6. 授業設計、実施担当者からの意見という6章編成となっている。入門科目3科目のうち、「eラーニング総論」の5章の一部を紹介する。

下に示した総合評価チャートは、受講申込者に対する「修了率」、1回でもCCSにアクセスした学生数に対する全ての教材を起動した学生数の割合を表す「教材使用度」、対面授業とeラーニング授業の比率に関する学生の評価である「eラーニング比率」、学習内容に対する学生の評価を示す「内容満足度」、同じくメンタへの評価を表す「メンタ」の5つの視点から正規授業とeLPCO授業の結果を表しています。



このチャートが示しているように、最大の問題はeLPCO授業における修了率の低さと、その原因となっている教材使用度の差です。（以下略）

フレーム1の評価の目標は、前期授業を担当した教員、インストラクショナルデザイナー、メンタ等への速報的なフィードバックを提供することであり、様々な角度からの授業改善に役立ててもらうことである。引用した部分以外にも、報告書には、学生がコンテンツにアクセスす

る時間帯のパターンやメンタへの質問傾向なども記述されており、実際の授業改善計画の基礎資料として用いられている。

5. フレーム2：データマイニングによる評価

この評価は、本プログラムの総体的な評価を示すモデル開発を目指している。各科目を一つの大きな上位変数と考え、その科目に属する情報を下位変数としまとめ、プログラムの評価を行い、より普遍的なeラーニングの特性を調査していこうという試みである。

分析方法としては、総体的評価のモデル作成のために、学習効果と満足度を従属変数にし、実施した科目から入手可能な情報を個々に変数化して統計処理、分析を行っている。類似の先行研究をみると、研究前にタイプIエラー、タイプIIエラーを統制するため、 $\alpha = .01$, $\beta = .2$, medium ES, .8 powerに設定しパワー分析を行っている(Cohen et al. 2003)。Cohenらの研究が示すように、変数は意義の似ている物をセットとして、まとめ分析し、統計的有意差の出たものに関してさらに詳細を確かめていく方法を取っている。

現在、具体的には学生の学習スタイルに応じた、効果的なeラーニングでの学習法と評価を結びつける研究を行なっている。

自学自習、自律的学習を可能にするeラーニングでは、学習者の多様性に対応した学習環境を提供することが重要になる。インタラクティブ性、協調学習の有無、学習者へどのように学習成果をフィードバックするのが適切であるのか等、各学習者に適合する教材、教授法等の研究を促進する必要がある。そこで、学習スタイルの違いに注目し、学習者の学習スタイル別の効果的、効率的な学習環境とは何かを明らかにするための研究を推進している。

本研究では、Felder & Silverman (1988) のモデルをベースに開発されたFelder & Soloman (2004) の学習スタイルインデックスを用いている。当該インデックスを採用した理由は、Felder & Solomanのモデルが、Riding & Rayner (1998) が多数の学習スタイルモデルを整理したCurryモデルのうち、最も外因の影響を受けにくく、生来の性格・気質・能力によるものとみなされる「認知・人格スタイル」という中核部に属している(青木2005)ためである。また、この学習スタイルインデックスに関しては、Felder & Spurlin (2005) において妥当性と信頼性についてまとめられており、研究結果を一般化しやすい点も採用理由の一つである。

授業ごとの形成的評価と総括的評価から得られたデータを用い、Felder-Silvermanモデルの4要因(外向的-内向的、感覚的-直感的、知覚的-判断的、順次的-全体的)との関係を調査している。Felder-Silvermanモデルでは、

学習スタイル別の効果的な学習方法への提案が簡単になされているが、この検証も同時に行っている。

評価フレーム2では、研究と教育実践を結ぶために、eラーニングの特性と学習の関係をより一般化できるような研究および評価活動を行っている。eLPCOは、青山学院大学総合研究所に設置され、教育実践だけでなく教育の場で役立つような研究活動を推奨されている。eラーニング利用による学習についての普遍的な傾向が明らかになれば、eラーニングの教育の質を向上させ、更なる教育実践への応用・普及が可能になると考え、それを目標に活動を進めている。

6. まとめと課題

本稿では、eLPCOのTF13が取り組んでいるeラーニングの質保証・評価フレームの形成過程、活動内容、2006年度前期に関する結果の一部を述べた。その中には、eラーニングのデータの多様性を反映して、多様な評価の対象や評価の方法が提案され、実施されていることを紹介した。

教材や学習プログラムの評価で重要なのは、評価結果を次の計画に活かすシステムであり、「評価から計画に戻る矢印、評価→計画」(鈴木 2002)である。単位を認定する正規授業としてのeLPCOのeラーニングプログラムは、2006年の4月に本格的に提供開始されたので、TF13の活動はやっとPlan-Do-Seeの1サイクルを終えたところであり、緒に就いたばかりであると言える。また、フレーム3、4の評価は2007年度以降に本格化する。

したがって、今後の課題はPlan-Do-Seeのサイクルを繰り返し、フレーム1評価の妥当性、信頼性を高めること、フレーム3、4評価のより具体的な評価項目を開発することである。また、フレーム2の研究によって新たな評価モデルが提案された場合には、その試行や位置付けにも積極的に取り組んでいきたい。

謝辞

本稿執筆にあたって、青山学院大学eラーニング人材育成研究センターTF13メンバの皆様にご協力をいただきました。深く感謝いたします。

参考文献

- 青木久美子, “学習スタイルの概念と理論及びそれに基づく測定方法：欧米の研究から日本独自の研究への考察”, NIME 研究報告71, 2005
- Asia e-Learning Network (AEN), “コンテンツ、サービスの品質基準等に関わる調査 (AEN-WG4)”, 2005
<http://www.asia-elearning.net/content/japan/act2004/index.html> [accessed on Dec. 12, 2006]
- Bourne, John and Janet C. Moore (eds.), “Elements of Quality

Online Education: Practice and Direction”, The Sloan Consortium, Needham MS, 2003

CEN/ISSS, “CWA Quality Assurance Standards”, CEN Workshop Agreements, 2003
<http://www.cenorm.be/cenorm/businessdomains/businessdomains/iss/cwa/learning+technologies.asp> [accessed on Oct. 8, 2004]

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S., “Applied Multiple Regression / Correlation Analysis for the Behavioral Sciences (3rd Ed.)”, Lawrence Erlbaum Associates, Inc., NJ, 2003

Felder, R. M. & Silverman, L. K., “Learning and Teaching Styles in Engineering Education”, Engineering Education, 78(7), pp.674-681, 1988

Felder, R. M. & Soloman, B. A., “Index of Learning Styles”, 2004
<http://www.ncsu.edu/felder-public/ILSpa.html> [accessed on Aug. 8, 2006]

Felder, R. M., & Spurlin, J., “Applications, Reliability and Validity of Index of Learning Styles”, International Journal of Engineering Education, 21(1), pp.103-112, 2005

Groccia, James E., “The Student as Customer versus the Student as Learner.” About Campus 2, 2, pp.31-32, 1997

Horton, William, “Evaluating E-Learning”, ASTD, Alexandria, VA, 2001

Kirkpatrick, D. L., “Evaluating Training Programs: The Four Levels. (2nd ed.)”, Berrett-Koehler Publishers, San Francisco, 1998

松田岳士, ‘メンタリングー授業実施戦略ー’ 玉木欽也監修 “eラーニング専門家のためのインストラクショナルデザイン”, 東京電機大学出版局, 東京, 2006

Mayadas, Frank A., John Bourne and Janet C. Moore, ‘Introduction’, In Bourne, John and Janet C. Moore (eds.), “Elements of Quality Online Education: Practice and Direction”, The Sloan Consortium, Needham MS, pp.7-10, 2003

North Central Association Commission on Institutions of Higher Education, “Best Practices for Electronically Offered Degree and Certificate Programs”, 2000
http://www.ncahigherlearningcommission.org/resources/electronic_degrees/ [accessed on Oct. 8, 2004]

Palumbo, David, “Learner Relationship Management: Impacting the Top and Bottom Line.”, Proceedings of ED-MEDIA 2003 Conference, 1, pp.130-131, 2003

Porter, Lynnette R, “Creating the Virtual Classroom: Distance Learning with the Internet”, John Wiley & Sons, Inc., New York, 1997

Riding, R., & Rayner, S., “Cognitive Styles and Learning Strategies: Understanding Style Differences in Learning and Behavior”, David Fulton, London, 1998

The Sloan Consortium (Sloan-C), “Effective Practices”, 2004
<http://www.sloan-c.org/effective/index.asp> [accessed on Nov. 10, 2004]

鈴木克明, “教材設計マニュアル”, 北大路書房, 京都, 2002

Tergan, Sigmar-Olad and Peter Schenkel, “Assessing the Instructional Power of E-Learning Applications: A Learner-Centered Checklist Approach”, Proceedings of ED-MEDIA 2003 Conference, 1, pp. 604-609, 2003

The Institute for Higher-Education Policy, “Quality on the Line: Benchmarks for Success in Internet-Based Distance Education”, The Institute for Higher-Education Policy, Washington D.C., 2000



まつだ たけし
松田 岳士
慶應義塾大学デジタルメディア・コンテンツ統合研究機構専任講師, 青山学院大学総合研究所客員研究員。1989年九州大学文学部卒。放送局勤務を経て, 青山学院大学国際政治経済学研究所博士後期課程修了, 博士(国際コミュニケーション)。コミュニケーションの視点から, eラーニングにおける学習活動の評価および学習者支援方法を研究している。日本教育工学会, 教育システム情報学会, AACE, AECT会員



こうだ よしこ
合田 美子
青山学院大学総合研究所客員研究員, 東京工業大学高等専門学校非常勤講師。1996年東京学芸大学大学院教育学研究科英語教育専攻修了, 筑波大学附属中学高校文部教官を経て, フルブライト奨学金にて留学, 2004年フロリダ工科大学科学教育博士(コンピュータ教育)取得。職歴は, 台湾樹徳科技大学応用言語学科専任, フロリダ工科大学大学院非常勤講師, NPO Marine Resources Council。現在の研究興味は, eラーニングにおける教育の質保証, 学習スタイルと学習活動の評価方法, コンピュータを使用した英語教育。日本教育工学会, 教育システム情報学会, AACE会員, JACET, KATE運営委員



たまき きんや
玉木 欽也
青山学院大学経営学部教授, 青山学院大学総合研究所eラーニング人材育成研究センター副センター長。1990年早稲田大学理工学研究科博士後期課程単位取得(工学博士)。主な研究テーマは, eラーニング人材育成, 事業創造戦略。日本教育工学会, 教育システム情報学会, 日本経営工学会, 経営情報学会会員。

Framework of Quality Assurance and Evaluation of e-Learning Using a Variety of Data

Takeshi Matsuda · Yoshiko Goda · Kinnya Tamaki

In this paper, the scheme of quality assurance of e-learning programs proposed by the Research Center for e-Learning Professional Competency or eLPCO is discussed in comparison with existing frameworks. Two major features of existing QA and program evaluation frameworks are emerged with the examination of some representative examples. First, most schemes assess e-learning programs on certain checklists and their objectives can be categorized into two evaluation systems; absolute evaluation and relative evaluation. Second, some exiting schemes adopt multiple levels of QA on the assumption that there should be classes of QA for complex processes such as e-learning.

Considering the examples of the existing frameworks, eLPCO decides to apply checklist approach in combination with data mining approach. Specifically, four frameworks of QA and evaluation are employed; 1 Evaluation of each course, 2 Creation of novel evaluation model as holistic evaluation, 3 Evaluation of certification program, and 4 Investigation of evaluation process. This paper describes the activities and progress of framework of 1 and 2.

Keywords

Frameworks of Quality Assurance, Checklist Approach, Data Mining Approach, Evaluation of Each e-Learning Program, Holistic Evaluation