

# 高等教育における学習成果アセスメント

## - 特筆すべき事例の比較研究 -

OECD 教育関連ワーキングペーパーNo. 15

Deborah Nusche 著

深堀 聡子(Satoko Fukahori) 訳

Assessment of Learning Outcomes in Higher Education: A Comparative Review of Selected Practices

OECD Education Working Paper No.15

©2008 OECD

原著（英語）は、OECDによって、上記のタイトルで刊行されたものである。  
この和訳版は、国立教育政策研究所が OECD パリ本部との協定にもとづいて刊行したものである。和訳の正確さおよび原著との統一性については、すべて国立教育政策研究所の側が責任を負っている。

©2008 National Institute for Educational Policy Research of Japan (NIER)

教育局 (DIRECTORATE FOR EDUCATION)

2008年2月20日づけの同文書を削除・差し替える

## 高等教育における学習成果アセスメント：特筆すべき事例の比較研究

OECD 教育関連ワーキングペーパーNo. 15

Deborah Nusche 著

### 著者について

デボラ・ニュッシュ氏は、2007年4月に、OECD 教育局の指標・分析課の顧問として、この論文を執筆した。ニュッシュ氏は現在、同局教育訓練政策課の政策研究員として、研究に従事している。

連絡先： 教育局 Deborah Nusche

[電話：+33 1 4524 7801 ; Email : Deborah.Nusche@oecd.org]

JT03241528

OLIS にて全文をオリジナル形式で閲覧可能

## OECD 教育局

### OECD 教育関連ワーキングペーパー・シリーズ OECD EDUCATION WORKING PAPERS SERIES

本シリーズは、OECD 教育局による優れた研究成果を幅広く公開することをめざしている。執筆は通常、共同作業によるが、主要な著者だけが明記されている。なお論文は原則として、もとの言語（英語またはフランス語）のみで公開し、他方の言語による要旨を添えている。

本シリーズに対するコメントを歓迎し、下記の連絡先で受け付けている。

- [edu.contact@oecd.org](mailto:edu.contact@oecd.org)
- Directorate for Education, 2, rue Andre Pascal, 75775, Paris CEDEX 16, France

各論文で示されている見解は、著者が全責任を負うものであり、OECD および加盟国政府の見解を必ずしも反映するものではない。

論文の全文、または一部の転載・翻訳許可の申請は、下記の連絡先で受け付けている。

- OECD Publishing,
  - [rights@oecd.org](mailto:rights@oecd.org)
  - FAX 33 1 45 24 99 30

## 要 旨

高等教育機関に、教育と学習の質に関するアカウントビリティ・データと消費者情報を提供するよう要請する社会的圧力は、ますます強くなってきている。ところが既存の高等教育機関の格づけやランキングは、学習成果に関する情報を見過ごし、活用された資源、開講された科目、刊行された論文といったインプット、活動、研究アウトプットに注目する傾向にある。それらの指標は、高等教育機関が学生の知識と技能の向上にいかに関与しているかについて、有効な情報を提供するものではない。高等教育機関やプログラムにおける教育の質に関する比較可能な情報が整備されている国は、ほとんど存在しないのである。

一部の国では、高等教育における学習成果を測定する方法が開発されてきているものの、各国で採用されている指標の特徴に関する情報は、ほとんど存在しない。本稿は、OECD 加盟国および連携国において、この領域について蓄積されてきた経験を整理したものである。18 のアセスメント・ツールについて記述し、そのデータにもとづいて、既存のアセスメント・ツールの概念的・組織的・方法論的特徴を整理する。そして高等教育における学習成果のタイポロジーを提示し、それぞれが各国においてどのように測定されてきたのかを明らかにする。注目する事例は、オーストラリア、ブラジル、カナダ、メキシコ、イギリス、アメリカにおけるものである。

## 目 次

要旨	3
はじめに	5
1. 本稿の範囲と限界	6
1.1 着目するアセスメント・ツールの種類	6
1.2 限界	7
1.3 情報源	7
2. 高等教育における学習成果のタイプロジー	8
2.1 学習成果の概念	8
2.2 いかなる学習成果を測定するのか	9
2.2.1 認知的成果	10
2.2.2 非認知的成果	12
3. 学習成果アセスメントの事例	16
3.1 何が測定されているのか	19
3.2 いかなる方法で測定されているのか	20
3.2.1 関与している組織とアセスメントの財源	20
3.2.2 アセスメントの頻度	25
3.3 アセスメントのデザイン	27
3.3.1 アセスメント・ツールの種類（技能アセスメント 対 自己報告）	27
3.3.2 形式	28
3.3.3 アセスメントにおける質の基準	29
3.4 対象者	32
3.4.1 ターゲット人口	32
3.4.2 サンプルの抽出	38
3.4.3 動機づけ	38
3.5 アセスメントの目的と結果	39
3.5.1 アセスメントの焦点	39
3.5.2 アセスメント結果の用途	41

参考文献(References) [省略]

注：情報源（国別）(Annex: Data Sources by Country) [省略]

## はじめに

高等教育機関に、教育と学習の質に関するアカウンタビリティ・データと消費者情報を提供するよう要請する社会的圧力は、この20年間に強まりを増してきた。学生の学習成果に関する比較可能な情報を要請しているのは、大学入学志願者、優れた卒業生を求める雇用者、公的資金の有効利用を求める納税者、適格認定や資源配分に関する意思決定に関わる政策立案者などのステークホルダーである。社会全体としても、知識基盤型の経済体制へと移行するなかで、高等教育機関がいかに効果的に学生を育成しているのかに関心を寄せている。高等教育機関も、教育改善、社会へのアカウンタビリティの遂行、消費者情報の提供にむけて、学生の学習成果を比較検討することから利益をえることができる。優れた学習成果の証拠は、公的資金を獲得したり、授業料を払う学生を引きつけたりするうえでも役立つであろう。

*USA News & World Report*や*Times Higher Education Supplement*をはじめとする、教育の質に関する既存の格づけやランキングは、学習成果<sup>1</sup>に関する情報を見過ごし、活用された資源、開講された科目、刊行された論文といったインプット、活動(activities)、研究アウトプットに注目する傾向にある。しかしながら、そうした指標では、高等教育機関が学生の知識と技能の育成にいかに貢献しているのかをとらえることはできない。したがって、これらの格づけやランキングは、政府、学生、社会全体に対して、教育と学習の質に関する情報を提供するツールとしては、不適切といわなければならない。ところが、高等教育機関の学習成果を比較可能にするアセスメント・ツールが存在しないため、教育の質を相対化するための代替指標として、格づけやランキングが広く利用されているのである。格づけやランキングは多くのメディアによって取り上げられ、高等教育機関や卒業生に対する社会認識に強い影響をおよぼしている。

多くの国々の政策立案者やステークホルダーによって、学生がそれぞれの高等教育機関で実際に何を学習しているのかを、比較可能な形で明らかにするツールを開発する必要があることが強調されている。一部の国では、高等教育における学習成果を測定して、機関ごとに比較する手法が既に開発されているものの、アセスメント・ツールの特徴に関する情報は国家間でほとんど共有されていない。

本稿のねらいは、各国で実践されている高等教育における標準化された学習成果アセスメントについて、国際的な観点から検討することである。第1節では、本稿の範囲と限界について整理する。第2節では、高等教育における多様な学習成果のタイポロジー

---

<sup>1</sup> 主要な国際ランキングで使用されている指標と方法論のレビューについては、Salmi and Saroyan (2006)、IHEP (2007)、Sadlak and Liu (2007)などを参照。

を整理し、各類型を学習の質の指標とみなすことの長所と短所について検討する。最後の第3節では、各国で多様な学習成果がどのように測定されてきたのかを明らかにする。

## 1 本稿の範囲と限界

本稿では、高等教育における多様な学習成果をタイポロジーとして整理し、学習成果を地域レベルあるいは国家レベルで測定するために設計された、18 のアセスメント・ツールの特徴を概説する。本稿で注目するのは、オーストラリア、ブラジル、カナダ、メキシコ、イギリス、アメリカの事例である。高等教育における学習成果の質を学科、課程、機関ごとに比較する際に、その結果を活用できるアセスメントの事例に限定して採用した。

### 1.1 着目するアセスメント・ツールの種類

本稿では何よりもまず、高等教育機関が学生の学習成果の促進にいかに関与しているのかに焦点をあてる、大規模で直接的なアセスメントに注目する。現在、高等教育段階のテストが全国的に実施されており、すべての高等教育機関の学生に受験が義務づけられている唯一の国は、ブラジルである。もっとも、大規模で直接的なアセスメントは、オーストラリア、メキシコ、アメリカでも実施されている。オーストラリアとメキシコでは、高等教育機関が卒業年度の学生を対象とした全国規模の標準テストに任意で参加することができる。アメリカでは、民間テスト業者によってさまざまなテストが提供されており、毎年何百もの高等教育機関によって利用されている。このように、悉皆調査ではないものの、幅広い高等教育機関によって利用されているアセスメント・ツールを用いて、課程や機関別に結果を比較することができる。

大規模な標準テストは、高等教育機関で広く実施されているわけではないため、本稿では、高等教育機関の学士課程段階での達成度の指標とみなすことのできる大学院入学試験にも注目する。大学院入学試験は、学生による大学院での研究を手掛けるために必要な知識と技能の習熟度を測定するものである。この種のテストは多くの国に存在する。本稿で取りあげるのは、二つの事例のみであるが、それは大学院入学試験の結果から高等教育機関の学生に関する結論を導くには、無視できない制約があるからである。すなわち大学院入学試験は、学士課程の教育目標と必ずしも整合性が確保されているわけではなく、大学院入学試験に向けて準備をする学生には、サンプルとして偏りも認められるからである (Klein et al., 2005)。

最後に本稿では、直接実施するテストだけではなく、学生の自己報告にもとづいて学

習成果を測定する事例にも注目する。オーストラリアの CEQ やアメリカの NSSE などの一部の調査では、学生に高等教育機関での経験が自らの学習成果にどれほどのインパクトがあったのかを評定するよう求めている。またイギリスの DLHE やオーストラリアの GDS は、労働市場における成果や進学に着目している。これらの調査から、学習成果の直接的な証拠をもたらすものではないが、高等教育の質の一部を明らかにする二次的な指標を含んでいる。

## 1.2 限界

本稿の目的は、既存のアセスメント・ツールに関する網羅的なレビューを行うことではない。本稿で提示する表は、各国で実施されている学習成果アセスメントについて、その多様な種類、設計、活用のあり方を明示するためのものである。

なお本稿では、各高等教育機関によって独自に手掛けられてきた、さまざまな学習成果アセスメントは取り上げない。外部機関によって開発・測定されているため、課程や機関別に比較可能な大規模な標準テストのみを検討の対象としており、その範囲を超えた議論を展開することは意図していない。

専門職団体または専門大学（法学大学院や医学大学院など）によって設計された職業資格試験やアセスメント・ツールも、本稿では検討の対象に含まない。

## 1.3 情報源

本稿の情報は、ウェブサイトより収集したものである。民間組織によって設計されたアセスメント・ツールに関する情報は、主として、テスト会社のホームページから入手した。政府機関によって手掛けられたアセスメント・ツールに関する情報は、教育省や統計局など、種々の公的機関のホームページから入手した。具体的なテスト実施の手順や結果の活用に関する情報は、アセスメントに関わる民間組織および個々の高等教育機関のホームページを参照した。情報収集に関する限界の一つとして、国家レベルのアセスメントに関するインターネット情報が、その国の言語でしか入手できない場合が多い点があげられる。本稿で紹介する事例は、英語、フランス語、スペイン語、ポルトガル語による情報にもとづいている。アセスメント・ツールの内容が更新・改定されている可能性があることにも留意する必要がある。本稿の情報は、すべて 2007 年 4 月に収集したものである。各テストに関する参考資料のリストは、巻末に掲載する。



## 2 高等教育における学習成果のタイポロジー

### 2.1 学習成果の概念

学習成果とは、学習の結果としてもたらされる個人の変化や利益を意味する。そうした変化や利益は、能力または達成度という形で測定することができる。オッター (Otter, 1992, p. i) によると、学習成果とは「学習者が学習の結果として知ようになった、またはできるようになったこと」と定義される。本稿では高等教育における学習成果に注目するため、個人的な成長や社会的成熟など、高等教育機関のコントロールのおよばない要因によるのではなく、高等教育の経験によると考えられる成果に焦点をあてる。すなわち本稿では、高等教育機関によって提供される学習機会を学生が享受した結果としてもたらされる学習成果について検討する。

学生の学習に焦点をあてるためには、インプット、活動(activities)、アウトプット<sup>2</sup>といった、教育の質をとらえるためにしばしば用いられる評価指標と、成果を明確に区別するアプローチをとる必要がある。これらの用語の誤用は少なからぬ混乱を招きうるため、一貫した定義を確立しなければならない。ここでインプットとは、使用された財政的、人的、物的資源をさす。たとえば、教育予算や寄付金、教員と事務局、建物や設備などが含まれる。活動とは、インプットを動員して特定のアウトプットを生み出すためにとられる行動や実施される作業である。高等教育における活動としては、たとえばカリキュラム編成や教育活動などがあげられる。アウトプットとは、機関や制度が輩出するあらゆるものをさす。高等教育機関におけるアウトプットは、刊行された論文、開講された授業、配布された教材、授与された学位といった形で測定することができる。インプット、活動、アウトプットは、学生による学習という観点からは、それ自体にほとんど価値はなく、成果や利益に結びつくかどうか不確定な媒介的要素にすぎない。

成果(アウトカム)は、教育機関が教えようと意図する内容ではなく、学生が実際に達成した内容をさす(Allan, 1996)。アイズナー(Eisner, 1979, p. 103)によると、成果とは「その意図にかかわらず、なんらかの取り組みの後に、最終的に身につけているもの」である。また多くの著者が、学習成果を観察、証明、測定することが可能なものと定義している(Spady, 1988; Melton, 1996)。教育機関が学習成果について言及する際には、その質の測定が可能であることが、しばしば示唆されている(Melton, 1996)。

学習成果という用語は、*成果を重視する教育 (outcomes-based education)* に由来し

<sup>2</sup> 以下で述べるインプット、活動、アウトプットの定義は、OECD Glossary for Results-based management (2000) で用いられている用語にもとづいている。

ている。成果を重視する教育とは、学生が何を学習すべきなのかを明確に識別、定義、測定することを特徴とする教育構造モデルである（Spady, 1988 ; Allan, 1996 ; Andrich, 2002 ; Adam, 2004）。成果を重視する教育システムでは、学生の学習成果が明確かつ詳細に定義され、それにもとづいてカリキュラムが編成される。ここで定義とは、学習者が一定の学習期間を終えたときに、どのような知識や理解に至っており、何ができるようになっていることが期待されているのかを明らかにするものである（Adam, 2004）。この成果を重視する教育のアプローチは、とくに中等教育段階の学校教育で頻繁に採用されている（Ewell, 2005）。高等教育においては、まずアメリカ、オーストラリア、ニュージーランド、イギリスで導入され、近年では他の OECD 加盟国でも手掛けられるようになってきている（Adam, 2004）。期待される成果という観点からカリキュラムを定義することは、学習成果アセスメントにむけた重要なステップといえよう。高等教育機関が、期待される学生の成果を明確かつ測定可能な形で特定したとき、学習成果アセスメントとその比較検討が可能になるのである。

もつとも、高等教育における成果は、学習成果に限定されない。学生は、高等教育機関での経験からさまざまな形で恩恵をうけることができる。より高い社会的地位、高い就職率、市民社会への参加、大学院への進学機会、あるいは単により充実した人生を送ることなどがその一例である。（Ewell, 2005）。しかしながら、そうした成果は学習と関連しているものの、高等教育機関における学習経験をとおして実際に習得された知識、能力、技能と混同すべきではない（Ewell, 2005）。高等教育の長期的、社会経済的利益を学習成果の二次的な代替指標とみなすことはできるが、直接的な学習成果とみなすことはできない。

## 2.2 いかなる学習成果を測定するのか

成果の意味を定義することは比較的容易であるものの、学習の範囲・内容をめぐるコンセンサスを形成することは容易ではない。学習には多くの要素があり、測定しやすいものも、測定しにくいものも含まれている。高等教育機関は、それぞれ異なるミッションと目標を追求しているため、教育と学習の優先事項も大きく異なっている。特定の専門分野別知識を伝授するのに適した課程もあれば、一般的技能やコンピテンスを教授するのに卓越した課程もある。一つのアセスメント・ツールで、高等教育の重要な成果をすべて包括的に測定することは、不可能といわなければならない。したがってアセスメントに携わる者は、自らの目的に対してもっともレリバンスが高く、アセスメントの対象となる高等教育機関のミッションと目標に対応する成果を選択するのがよいだろう。

本節では、高等教育における学習成果のタイポロジーを提示し、高等教育におけるさまざまな学習成果を定義、記述、分類することをめざす。そして各類型の成果を、高等

教育における学習の質の指標とみなすことの長所と短所について検討する。タイポロジーは、認知的なおよび非認知的な学習成果といった古典的分類にもとづいているが、「コンピテンス」の概念も導入し、既存のアセスメントの取り組みの多くが依拠している、認知的・非認知的な二分法を乗り越えることを試みる。

もっとも学習を構成する諸要素は相互に関連、重複しているため、いかなる学習成果の分類も、一定の作為性を伴うことは避けられない点に留意する必要がある（Bowen, 1977）。ここで提示するタイポロジーは、学習とそれを測定するために用いられる学力テストのさまざまな特徴をとらえるための実践的枠組みにすぎない。

### 2.2.1 認知的成果

本研究で取り上げるアセスメント・ツールは、そのほとんどが認知的な学習成果の測定に焦点をあてている。認知的な学習とは、知識の回想や知覚、認知的能力や技能の発達を意味する（Posner, 1992）。そして広義の認知的な学習成果とは、「専門分野別の知識から、極めて一般的な論理づけ能力や問題解決能力を含む幅広いもの」（Shavelson and Huang, 2003, p. 13）として理解される。

認知的な学習成果の類型は、数多く存在する（Gagné, 1977 ; Kolb, 1981 ; Klemp, 1977 ; Eraut, 1990 ; Marzano, 2001）。その多くはブルーム（Bloom, 1956）による教育目標の分類学（タキソノミー）の影響を受けている。このブルーム・タキソノミーは、教育をとおして育成しようとする成果の枠組みを示すものであり、知識の理解だけでなく、応用、統合、分析、評価などの学業的スキルも含むものである。ブルーム・タキソノミーは、今日も基本的枠組みとみなされているものの、多くの著者によって批判され、展開され、修正されてきた（Anderson and Sosniak, 1994）。

本稿の目的に照らし合わせると、ここでは成果を知識とスキルに分類するだけで十分といえよう。この基本的分類は、アセスメントに関する先行研究で広く採用されてきた。

#### 知識成果 (*Knowledge Outcomes*)

知識の習得とは「概念、事物、事象について、認知または回想によって記憶すること」（Bloom, 1956, p. 62）を意味する。知識成果のアセスメントには、一般的内容の知識に注目するものもあれば、専門分野別の知識に注目するものもある。

一般的知識とは、「必須の学習内容」（Maeroff, 2006）とみなされているコア・カリキュラムを構成する知識をさす。たとえば、ブラジルで学生の学力を測定するために実施

されている全国テスト（ENADE）には、社会的・人種的多様性、公共政策、国内外の時事問題などの一般的知識を問う項目が含まれている（Verhine and Dantas, 2005）。一部のステークホルダーは、すべての学生が専攻に関わらず、一連の必須知識を習得すべきだと主張しており（AAC&U, 2004）、一部の国では、そうしたコア・カリキュラムが学士課程教育において重要な位置を占めている。一般的知識を学科、課程、機関別に測定することは、すべての学生にとって必須とみなされている「共通」の学習内容の推進に、各機関がどの程度に成功しているのかを比較検討するうえで、有効といえる。もっとも、一般的知識は、高等教育プログラムで教えられている内容のほんの一部分にすぎないため、一般教育の成果を教育の質の唯一の指標とみなすのは、不十分といえよう。

*専門分野別・科目別の知識*とは、生物学や文学など、特定の専門分野で習得される知識をさす。専門分野別知識の成果に着目したアセスメントは、特定の分野における学習の質を機関間で比較する場合にとくに有効である。ある専門分野に特化した科目履修の機会は、疑いなく高等教育のもっとも重要な部分であり、高等教育機関も学生を各専攻分野において重点的に指導するよう義務づけられている（Volkwein, 2003）。パスカレラとテレンジニーニ（Pascarella and Terenzini, 2005）によると、学生は高等教育機関在学中に、自己の専攻と合致する分野において、もっとも著しい成長を遂げる。さらに専門分野別の知識の習得は、他の種類の成果と比較して、高等教育機関で教育・学習されている内容と、はるかに強く結びついている（Allan, 1996）。したがって高等教育機関のアカウンタビリティは、学生がそれぞれの専攻分野においてどれだけ専門的知識を習得したかという点について、もっとも重点的に問われるべきであろう。

### *技能成果 (Skills Outcomes)*

認知的な*技能*は、言語的・数量的論理づけ、情報処理、理解、分析、批判的思考、問題解決、新しい考え方への評価といった、複雑な思考プロセスから成り立っている。そうした思考プロセスが、一般的な（一般化されたパターンに従う）ものなのか、専門分野別に固有のものなのかをめぐって、見解は一致していない。

学習成果の学科別比較を試みるアセスメントでは、*一般的技能の成果*に着目することが多い。すべての一般的技能に共通する特徴は、学問分野を横断して確認される点にある。一般的技能は、どのような科目や状況においても援用することができ、特定の学科に直接結びついているものではない。あらゆる学問分野に通用し、新しい状況下でも学生が活用できるものである（Pascarella and Terenzini, 2005）。一般的技能の成果は、知識よりも応用を重視し、学生の知的問題を解決する能力に着目するテストを用いて測定することができる。通常、学生は自ら論旨を組み立てて解答を作成するよう求められるが、それによって文章記述力も試される。技能という観点から成果に着目することに

よって、著しく異なるミッションや教授法をもつ課程や機関が、ある共通の技能を学生に習得させることにむけて、どれほど成功しているのかを比較検討することができるのである。

もつとも、そうした成果が大学での経験に起因するのかは、必ずしも定かではない。高等教育機関の多くが、全学的なミッションや目標のなかで一般的技能の育成を強調しているものの、特定の学科のカリキュラムのなかに明確に位置づけられることはほとんどない。それでは、一般的技能は実際にどこで習得され、大学はその育成にいかなる役割を果たしているのだろうか。学術的技能という観点からみた学生の発達、学生の年齢と相関をもつことを明らかにした研究もある (Banta and Pike, in Ewell, 1991)。そうした年齢と能力との相関関係から示唆されるのは、技能の成果の一部は、高等教育機関での学習ではなく、社会的成熟の結果としてもたらされることである (Ewell, 1991)。したがって、一般的技能のアセスメントは、高等教育機関による学生の学習へのインパクトよりも、学生の知能や過去の学校教育の結果を主に測定する結果に陥る危険性をはらんでいる。

**専門分野別技能**とは、自然科学や人文といった、学問分野の大枠のなかで共用されている思考パターンをさす。特定の学問分野における探究の方法、証拠を検証する方法、新しい状況に対処するためとるべき手続きのパターン等として定義されている。特定の知識をどのように、なぜ、いつ適用するのかを理解していることが前提となる (Shavelson and Huang, 2003)。分野別技能は、必ずしも科目を越えて援用することはできない。たとえば、「卓越した文章記述能力」という能力成果は、学問分野によって異なる形式をとり、異なる技能を必要とする (AAC&U, 2004)。学業的技能は高等教育の一般的な成果の一つとみなされているが、学問分野の内容と切り離して検証することはできない。したがって分野別技能は、各学生に専門分野から選んだ新規の文書を与え、その論旨の質を評価させ、それにもとづいて複雑な解答を記述させることで測定することができよう (AAC&U, 2004)。

### 2.2.2 非認知的成果

非認知的な発達とは、信念の変化や、特定の価値観の形成をさす (Ewell, 2005)。高等教育機関のミッションには、しばしば非認知的な要素が含まれているが、これは高等教育機関の役割が知識と技能を習得させることにとどまらないことを示している。非認知的な成果は、教科指導のほか、高等教育機関がカリキュラムを補足するために企画する正課外の活動をとおして形成される。こうした「正課併行 (co-curricular)」の活動には、個別指導、チューター制度、カウンセリング、学生-教員交流、クラブ、運動競技などの活動が含まれる。このようなサービスの存在は、高等教育機関が、教科指導を

とおした学習を補足する手段として、非認知的な学習の発展にも価値をおいていることを明らかにしている (Middle States Commission on Higher Education, 1996)。

高等教育における非認知的な成果を理論的ないし経験的に類型化する試みは、少なからず実施されてきた。非認知的な学習成果に関する研究の多くは、アイデンティティ形成における特定の理論的段階が確認されるかどうか注目する (Pascarella and Terenzini, 2005)。パスカレーラとテレンジニーニ (Pascarella and Terenzini, 2005) は、アメリカの大学が学生におよぼすインパクトに関する 2,500 以上の調査に注目し、総括的に分析している。とりわけ多くの研究で測定されているのは、心理社会的 (psychosocial) 発達、態度、価値観に関連する成果であった。

心理社会的な発達には、アイデンティティ形成や自尊心などの自己形成のほか、生徒の他者、組織、環境などとの関係性の形成も含まれる。関係性の成果としては、対人関係スキルや異文化間スキルに加えて、自律性や成熟などがあげられる。態度と価値観は、互いに密接に関連しており、しばしば区別がつきにくい。両者を区別するならば、態度が特定の対象に関する信念であるのに対して、価値観は態度を越えて一般化された基準とみることができる (Rokeach, in Pascarella and Terenzini, 2005)。態度と価値観の成果としては、社会的責任、学習意欲、多様性への理解などがあげられる (Volkwein, 2003)。

アスティン (Astin, 1984) によると、学習成果は、単に高等教育機関における教育の質の産物というよりも、むしろ高等教育機関が提供する学習機会に、学生がいかに積極的に参加したかに規定されるものである。この考え方にもとづいて、NSSE (National Survey for Student Engagement, アメリカ) では、高等教育機関が学生の積極的な取り組みをいかに活発に促しているのかを測定することがめざされている。

アセスメントに向けて、非認知的な学習成果の何を望ましい成果と定義するのかをめぐって、必ずしも合意は得られていない。いかなる態度や価値観を望ましく、「正しい」ものとみなすのかについての判断は、すべてのステークホルダーによって常に共有されているわけではない。しかも望ましい非認知的成果をめぐる定義は、文化的コンテキストによってかなり大きく異なることが予想される。さらに、非認知的な発達が真に大学での経験に起因するのかが疑わしい場合もある。すべての学生が大学の課外活動に参加しているわけではなく、正課外の経験のすべてが大学生活に関係しているわけでもない。非認知的な成果は社会的成熟や世代効果 (generational effects) (Pascarella and Terenzini, 2005)、ないし「重大なライフ・イベント (significant life events)」 (Glenn in Pascarella and Terenzini, 2005, p. 272) によることを示す先行研究も存在する。学生の発達におよぼす高等教育機関のインパクトを測定する際、これらの要因を統制する

ことは極めて困難といえる。

高等教育における非認知的な成果の研究は、認知的な成果の研究よりも複雑である。価値観や信念と、観察可能な活動や行動との関連は未だ明らかにされていない (Pascarella and Terenzini, 2005)。価値観や信念を行動データによって捉え、測定することが可能であるという証拠は、ほとんど存在しないのである。それゆえ非認知的な成果は通常、学生による自己報告や教員や雇用者を対象とする質問紙調査によって間接的に測定されている。そのような間接的な指標は、事実だけでなく、個人による物事の見方や感じ方に規定される。それゆえその測定結果は、知識や技能を直接測定した場合と比較すると、学生による学習の指標として、客観性に劣ると考えられる。

成果の認知的・非認知的分類を越えて、ライシャ (Rycher, 2004, p. 7) は、学習成果のより一般化されたモデルを提唱している。すなわち「コンピテンスとは、要求に応え、課題を的確に遂行する能力であり、*認知的・非認知的側面の両方を含む*」ものである。ウィンチとフォーマンペック (Winch & Foreman-Peck, 2004, p. 4) によると「コンピテンスは、不断に変化する状況のなかで、行動、知識、価値観、目標を組み合わせることから生まれる」。成果に関する既存のタイポロジーでは、多様な技能を組み合わせることによって真の専門性をもたらす成果の統合性を適切にとらえることができないという考え方に、学生の成果に関するこれらの概念はもとづいている (Ewell, 2005)。この考え方によると、生徒の成果は実践的課題のなかで定義され観察される必要がある。

コンピテンスに着目するアセスメント・ツールは、認知的、情緒的、行動的特性が複雑に組み合わさった状態を捉えようとする。たとえば学生のポートフォリオには、筆記課題、実習、インタビュー、実験やインターンシップの報告書といった、学生の取り組みの直接的な証拠が含まれる (Ewell, 2005)。さらにポートフォリオには、成果の間接的な証拠、すなわち学生にコンピテンスという観点から自己の成長を評価させる質問紙などが含まれることもある。学生による作品を多元的に評価することで、広範囲の学習成果を統合することが可能になる (Ewell, 2005)。

オター (Otter, 1992) は、さまざまな状況に適用される広範な能力としての一般的コンピテンスと、特定の職業分野で働くために必要なより狭い範囲の能力としての職業的コンピテンスの違いを強調している。職業的コンピテンスは、エンプロイアビリティと呼ばれることもある。学生が職場で役立つように準備することは、高等教育の主要な目的の一つといえる (Bowen, 1977)。高等教育機関が、産業、商業、その他のサービスのニーズに応える形で学生の能力を育成することに対する社会的要請は、ますます強くなっている (Seagraves et al., 1996)。雇用の際に高く評価される技能の育成も、そこに含まれている (Dearing, 1997)。

しかしながら、職業的コンピテンスという点に注目して学習成果を定義することには、いくつか留意すべき問題点もある。高等教育機関がもっぱら職業的コンピテンスを重視した場合、深く探究する機会を提供したり、学生の潜在能力の開発を支援したりといった、高等教育の重要な側面が軽視されるおそれがある (Otter, 1992 ; Melton, 1996)。また、卒業生が就く職業の領域は広範であるため、それぞれの学問分野について、明確な職業目標を定義することは必ずしも可能ではない (Melton, 1996)。学問分野によっては、職業的な役割を明確に定義できない場合もある。さらに職業的成果に関する事項は、即時的な雇用のニーズに焦点をあてる傾向にあるが、学生は初職のみならず、生涯にわたり不透明な労働市場で役立つ知的技能の習得により関心をもつと予想することもできる (Melton, 1996 ; AAC&U, 2004)。

望ましい職業的コンピテンスを国家資格制度のなかで定義している国もある。たとえばイギリスでは、高等職業資格は高等教育機関で単独に、あるいは他の学位と組み合わせで授与されている (Morgan, 2002)。イギリスの資格・カリキュラム機構 (Qualifications and Curriculum Authority, QCA) によると、そうした高等職業資格は「一定の基準を満たす技能、知識や理解のレリバンス、課題に対して適切に技術を活用したり知識や理解を応用したりする能力を測定することによって、・・・職場におけるレリバンスが明らかに高く、就職、昇進、教育訓練の継続に役立つコンピテンスを有していることを証明するものである」(QCA ホームページ、2007)。職業的コンピテンスを測定する一つの方法として、候補者の職場における能力を反映するポートフォリオの作成があげられる。その中身には、候補者の観察、活動の録音・録画、作品などが含まれる (QCA ホームページ、2007)。

就職率や卒業生の進路調査も、高等教育機関が学生の職業的コンピテンスをいかに高めているのかを測定する方法として、しばしば採用されている。卒業生調査からは、学生が高等教育機関における教育を、職業的成果の観点からどれほど有益と感じているのかを知るうえで、重要な情報を含んでいる。

もっとも、労働市場の成果は、必ずしも実際に習得されたコンピテンスを正確に反映しているわけではない点に留意する必要がある。就職率は高等教育の学習成果だけではなく、社会経済的要因やそれ以前の学校教育 (Bratti et al., 2001)、専攻分野 (McGuinness, 2003)、社会的ネットワーク (Montgomery, 1991)、高等教育機関のコスト (Pacarella et al., 1992) といったさまざまな要因に規定されていることが、多くの先行研究より示されている。スペンス (Spence, 1973) による労働市場におけるシグナリング理論によると、雇用者は選抜性の高い高等教育機関の卒業生の採用を好むが、それは能力の低い学生よりも教育コストが安価な有能な学生が、選抜性の高い高等教育機



関により高い比率で在籍しているという考え方にもとづいている。この理論によると雇用者は、高等教育機関の質よりも、その選別配分機能と応募者の能力に関する情報を提供する機能を重視していることになる。

### 3 学習成果アセスメントの事例

本節では、特筆すべき学習成果アセスメントの事例について概説する。すなわち高等教育における学習成果を、地域レベルまたは国家レベルで測定することを目的とする18のアセスメント・ツールの特徴を整理する。事例はオーストラリア、ブラジル、カナダ、メキシコ、イギリス、アメリカから採用した。本節では、これらの具体的事例について比較検討し、学習成果を測定するにあたっての概念的、方法論的、制度的側面に関する情報を提示する。本節では次の4つの課題について明らかにすることをめざす。

1. *何が測定されているのか。* セクション3.1では、各アセスメント・ツールで測定されている成果の種類に関する情報を整理する。
2. *それらの成果はいかなる方法で測定されているか。* セクション3.2では、テストの設計、作成、実施、報告の手順を明らかにする。セクション3.3では、アセスメント・ツールの形式、項目数、期間といったテクニカルな側面について整理する。
3. *アセスメント・ツールの対象者はだれか。* セクション3.4では、アセスメントの対象者の特徴を概説し、ターゲット人口、サンプルの抽出法、テストの適用範囲、学生や機関に対するアセスメントへの参加にむけた動機づけの措置について整理する。
4. *何のためにアセスメントが実施されているのか。* セクション3.5では、アセスメントで個々の学生、課程、機関、教育制度のいずれを測定しようとしているのかを明らかにする。どのような結果がめざされており、それがさまざまなステークホルダーによってどのように利用されるのかについて概説する。

各セクションで示すさまざまな特徴は、相互に密接に結びついていることが、全体をとおして明らかになるであろう。たとえば、いかなる調査デザインを選択するかは、とらえようとする成果の種類と、結果の使用目的に大きく規定されている。

表 1. 本研究で注目するアセスメント・ツールで測定されている成果

国	テスト名	開始年	測定されている成果の種類
オーストラリア	GSA (Graduate Skills Assessment)	2000	一般的技能：批判的思考力、問題解決能力、文章表現力（ACERは現在、基礎能力、マネジメント能力、ITスキル、リサーチ能力などを追加する方向で修正を検討中）。 <b>専門分野別知識と技能</b> ：（現段階では含まれていないが、ACERは大枠としての学問領域別にその要素をテストする方向で検討中）、 <b>非認知的成果</b> ：対人理解。
	CEQ (Course Experience Questionnaire, Graduate Destination Surveyの一部)	1993	一般的技能：問題解決能力、分析力、文章表現力。 <b>非認知的成果</b> ：チームワーク能力、次の項目に関する学生の満足度（指導、目標と水準、課題の量、評価）。 <b>一般的コンピテンス</b> ：未知の状況に対処する自信、仕事を計画する能力。
	GDS (Graduate Destination Survey)	1972	<b>職業的コンピテンス</b> ：卒業後約4か月目の就業状況（雇用の機会、雇用の職種、平均年収、就職活動）、進学状況（就学形態 - フルタイム・パートタイム、就学段階、学問分野、機関）。
ブラジル	ENC/Prova (Exame Nacional de Cursos)	1995-2003	<b>専門分野別の知識と技能</b> ：専門分野ごとにすべての高等教育カリキュラムに共通しており、必須とみなされている知識と技能。26分野について準備されている。
	ENADE (Exame Nacional de Desempenho dos Estudantes)	2004	<b>専門分野別の知識と技能</b> ：専門分野ごとにすべての高等教育カリキュラムに共通しており、必須とみなされている知識と技能。13分野について準備されている。 <b>一般的知識</b> ：テストで取り上げられてきたテーマ（生物学的・社会的多様性、公共政策、社会ネットワーク、市民性、時事問題）。 <b>一般的技能</b> ：推論、韻文の解釈、共通点を見出す、関係性の識別、省察、演繹、図表の解釈。
カナダ	NGS (National Graduate Survey, Follow-up Survey of Graduates)	1987	<b>職業的コンピテンス</b> ：卒業後2年目と5年目の就業状況（卒業後就いたすべての職業の数、特徴、期間、就職活動の期間、教育と職業の対応）、高等教育機関での経験に対する満足度。
	YITS (Youth in Transition Survey)	—	<b>職業的コンピテンス</b> ：第3回調査（調査時点でのターゲット人口の年齢22～24歳）に中等後教育とそれへの取り組み、就業状況に関する質問項目が含まれている。
メキシコ	EXANI-III (Examen Nacional de Ingreso al Posgrado)	1997	一般的技能：言語的・数学的論理づけ能力、推論・分析・総括する能力、情報処理能力（情報の整理・収集・理解）。
	EGEL (Examen General Para el Egreso de la Licenciatura)	1994	<b>専門分野別知識と技能</b> ：専門分野ごとにすべての高等教育カリキュラムに共通しており、必須とみなされている知識・技能。33分野について準備されている。
	EGETSU (Exámenes Generales Para el Egreso del Técnico Superior Universitario)	2000	<b>専門分野別知識と技能</b> ：学生の専攻分野で必要とされる理解力と問題解決能力。すべての工学系分野(19)について準備されている。一般知識と一般技能：社会・経済的知識、IT、英語など、すべての職業分野で必要とされる知識と技能。

英国	DLHE (Destinations of Leavers from Higher Education, First Destination Supplement に代わる)	2002	職業的コンピテンス：卒業後6ヵ月目の就労・進学状況（就業した卒業生の人数、職種、進学した卒業生の人数）。
	CAAP (Collegiate Assessment of Academic Proficiency)	1988	一般的技能：記述力（客観的・小論文）、読解力、数学力、科学的論理づけ能力、批判的思考力、すべての学問領域から取り込んだ教科内容。
	MAPP (Measure of Academic Proficiency and Progress, ETS Academic Profileに代わる)	2006 (1992-2006)	一般のおよび専門分野別技能：人文・社会科学・自然科学のコンテキストのなかで読解力と批判的思考力を測る。記述力、数学力。
アメリカ	Tasks in Critical Thinking (廃止)	1992	一般のおよび専門分野別技能：大枠としての学問領域（人文・社会科学・自然科学・芸術）のコンテキストのなかで、実践的課題が出される。課題の設定、分析、コミュニケーション能力。
	Major Field Tests (GRE Subject Test)	1990	専門分野別知識と技能：専門分野ごとにもっとも重要とみなされている知識と技能（知識、問題分析・解決能力、関係性をとらえる力、専門分野に関する図表などの資料を解釈する力）15の学士課程の学問分野と経営学修士について準備されている。
	CLA (Collegiate Learning Assessment)	2002	一般のおよび専門分野別技能：批判的思考力、分析的論理づけ能力、文章表現力、情報活用力。コンピテンス：実社会における課題（文書やデータを理解、評価、活用してレポートや政策提言を構築する）、大枠としての学問領域（人文・社会科学・自然科学・芸術）のコンテキストのなかで設定された実践的作業。
アメリカ	NSSE (National Survey of Student Engagement)	米2000加2004	非認知的成果：学生による取り組みに関する情報（学部学生の時間の使い方、授業から学んだ事から、課外活動、高等教育機関によるサービス）。
カナダ	GRE (Graduate Record Examination, General Test)	1966	一般的技能：言語的論理づけ能力、数量的論理づけ能力、分析的論述。

### 3.1 何が測定されているのか

前のセクションで示したタイポロジーにもとづいて、各アセスメント・ツールがとらえようとしている成果の種類を整理したのが、表1である。

学習成果の直接的なアセスメントのほとんどが、認知的な学習成果に焦点をあてている。ブラジルの ENC-Provao とその改訂版 ENADE では、すべての高等教育機関が専門分野ごとに共有すべき必須カリキュラムとみなされている、専門分野別の知識と技能に焦点があてられている。ENADE では、一般的な知識と技能の測定にあわせて、13 の専門分野について個別のテストが実施されている。

メキシコで実施されているテストでも、同様のアプローチがとられている。すなわち一般の卒業試験 (EGEL) では、33 の学問分野における専門分野別に必須の知識と技能に焦点があてられている。技術系大学の卒業試験 (EGETSU) は、19 の専攻分野に対応している。もっとも、EGETSU には英語や IT をはじめとする、ほとんどの分野で必要とされる一般的技能のアセスメントも含まれている。

アメリカで使用されているアセスメント・ツールの多くでも同様に、一般のおよび専門分野別の認知的な学習成果が測定されている。たとえば、MFT (Major Field Test) は、15 の学士課程分野と経営学修士 (MBA) に対応しており、CLA (Collegiate Learning Assessment) は、大卒としての学問分野 (自然科学、社会科学、人文、芸術) における実践的課題に焦点をあてている。

アセスメント・ツールのなかには、主に一般的な知識と技能に注目し、専門分野別の項目を含まないものもある。オーストラリアの GSA (Graduate Skills Assessment) と CEQ (Course Experience Questionnaire) は、批判的思考力や、問題解決能力、文章表現力などの技能に着目しており、ACER はさらに基礎的マネジメント能力、IT スキル、リサーチ能力などを追加することも検討している。メキシコの EXANI-III、アメリカの GRE をはじめとする大学院入学試験では、言語的・数学的論理づけ能力や情報処理能力などの一般的技能が測定されている。

学生に対する調査や質問紙は、非認知的な成果やコンピテンスに焦点をあてるものが多い。オーストラリアの CEQ では、卒業生に自らのチームワーク能力、計画的に作業を進める能力、未知の状況に対処する自信などを評価させている。アメリカおよびカナダで使われている NSSE (National Survey of Student Engagement) では、学生の高等教育機関での経験への関わり方に焦点があてられている。たとえば学生の時間の過ごし方、授業から学んだ事から、高等教育機関のサービスの利用状況などが調査の対象に含まれ

ている。

またオーストラリアの GDS (Graduate Destination Survey)、カナダの YTS (Youth in Transition Survey)、イギリスの DLHE (Destinations of Leavers from Higher Education) などの卒業生を対象とする調査では、近年卒業した学生がどれほど職業的コンピテンスを身につけているのかが明らかにされている。通常、卒業生の就職先や進学の有無に関する情報が提供されている。もっとも、労働市場での成果は、実際に習得されたコンピテンスを正確に反映するとは限らず、卒業生の職業的コンピテンス以外の多くの要因の影響をうけていることは、第 2 節で述べたとおりである。

### 3.2 いかなる方法で測定されているのか

このセクションでは、表 2 にもとづいて、本研究で検討する 18 のアセスメント・ツールの基礎的な情報を整理する。すなわち各アセスメント・ツールの財源、設計、実施、運用について明らかにする。さらにこのセクションでは、アセスメントの過程にどのような組織が関与しており、アセスメントがどのような頻度で実施され、結果がどのように報告されているのかについて比較検討する。

#### 3.2.1 関与している組織とアセスメントの財源

表 2 に示すとおり、検討するアセスメント・ツールのすべてが、アセスメントの対象となる高等教育機関の外部組織によって設計されている。通常、外部の業者が実施済みのテストの処理をして、高等教育機関に報告書を作成している。

アセスメントに関与している組織には、政府機関や、全国規模の統計機関、高等教育機関の協会、民間テスト業者、非営利団体などが含まれる。EGEL や EGETSU (メキシコ) をはじめとする一部の全国テストでは、学術代表者がテスト開発に携わっている。

民間テスト業者によって開発されたテストがもっとも広く利用されているのは、アメリカである。高等教育機関は直接テスト業者にアセスメント資材を発注し、キャンパスでテストを実施している。アメリカでもっとも有力なテスト業者は ETS と ACT である。ブラジルの全国テスト (ENADE) も民間テスト業者によって開発・分析されているが、公費で賄われており、高等教育機関がコストを負担することはない。オーストラリアでは、GSA の受験枠の一部が教育科学訓練省 (Department of Education, Sciences, and Training) によって負担されているが、追加の受験枠については、高等教育機関が負担しなければならない。メキシコでは学生自身が、全国卒業試験 (EGEL) の受験料を支払わなければならない。

表2. アセスメントの方法

国	テスト名	開始年	実施主体・テスト開発の スポンサー	テスト設計・分析の 責任者	実施形態	実施頻度	一時点・クロスセクショナル・長期的アセスメント	アセスメント結果の報告
オーストラリア	GSA (Graduate Skills Assessment)	2000	連邦政府：教育科学訓練省DEST (元 DETYA)	ACER (Australian Council for Educational Research)	高等教育機関で実施、教員が監督	年2回 (入学時・卒業時)	クロスセクショナル (入学時・卒業時)	学生に各要素の得点と同年受験者全体にしめる相対的位置づけを個別に報告。高等教育機関に学生の成績データ、機関単位の集計結果を報告。
	CEQ (Course Experience Questionnaire, Graduate Destination Surveyの一部)	1993	連邦政府：教育科学訓練省DEST (元 DETYA)	GCCA (Graduate Careers Council of Australia); ACER	GCCAが質問紙を高等教育機関に供給、高等教育機関が学生に送付	年1度	一時点	DESTへ各高等教育機関のデータを報告。高等教育機関へ概要を報告。大学、GCCA、ACER、政府によって多様な集計結果が公開される。報道機関および商業出版“The Good Universities Guide”によって、一般向けにランキングが発表される。
	GDS (Graduate Destination Survey)	1972	連邦政府：教育科学訓練省DEST (元 DETYA)	GCCA; ACER; ITS (University of Melbourne's Information Technology Service)	GCCAが質問紙を高等教育機関に供給、高等教育機関が学生に送付	年1度	一時点	DESTへ各高等教育機関のデータを報告。高等教育機関へ概要を報告。大学、GCCA、ACER、政府によって多様な集計結果が公開される。全国集計、プレス報道、GradStats (4頁綴りの結果) が発表される。
ブラジル	ENC/Prova (Exame Nacional de Cursos)	1995-2003	連邦政府	CESGRANRIO (テスト専門機関)	INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anisio Teixeira)の指揮のもとに、地方自治体によって、全国一斉に実施	年1度	一時点	結果はテクニカル・レポート、高等教育機関報告書 (専門分野・専攻別)、学生報告書の形で開示。学生報告書は本人にのみ提供。毎年専攻別格づけを公表。メディアによるランキングが行われる。
	ENADE (Exame Nacional de Desempenho dos Estudantes)	2004	連邦政府	CESGRANRIO, CESPE (テスト専門機関)	INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anisio Teixeira)の指揮のもとに、地方自治体によって、全国一斉に実施	3年に1度	クロスセクショナル (入学生・卒業年度生)	結果はテクニカル・レポート、高等教育機関報告書 (専門分野・専攻別)、学生報告書の形で開示。学生報告書は本人にのみ提供。結果は個別に報告されており、メディアの注目をあびることはない。

カナダ	NGS (National Graduate Survey, Follow-up Survey of Graduates)	1978	連邦政府 : Human Resources Development Canada (HRDC)	Statistics Canada (StatCan for HRDC)	StatCan地方支部による電話インタビュー (Computer Assisted Telephone Interviews, CATI)	定期的に実施。NGSを1978, 1984, 1988, 1992, 1997, 2002年に実施し、1987年以降は同じ回答者を3年後に追跡	同じ回答者の長期的調査 (卒業後2年目と5年目)	個別学生および高等教育機関の情報は開示されていない。高等教育機関は機関用結果概要を入手することができる。公開データでは、高等教育機関タイプと地域のみ識別可能。
	YITS (Youth in Transition Survey)	2000	連邦政府 : Human Resources Development Canada (HRDC)	Statistics Canada (StatCan for HRDC)	StatCanによる電話インタビュー (Computer Assisted Telephone Interviews, CATI)	2年に1度	長期的調査 (15歳&18~20歳コーホート)	集計結果を公開。公開データで個人は識別不可能。
メキシコ	EXANI-III (Examen Nacional de Ingreso al Posgrado)	1997	ANUIES (Asociacion Nacional de Universidades e Instituciones de Educacion Superior)	CENEVAL (Centro Nacional de Evaluacion para la Educacion Superior)	CENEVALによって高等教育機関のキャンパスで実施	所定の日時に実施される全国テストに、学生または高等教育機関が申し込む	一時点	学生個人のデータがインターネットをとおして、学生本人および学生の受験した機関に対して閲覧可能になる。
	EGEL (Examen General Para el Egreso de la Licenciatura)	1994	ANUIES	CENEVAL	CENEVALによって高等教育機関のキャンパスで実施	所定の日時に実施される全国テストに、学生または高等教育機関が申し込む	一時点	結果は非公開。高等教育機関は機関用結果概要を受け取る。学生は自己の得点を示す報告書を受け取る。全国平均以上を得点した学生を表彰。
	EGETSU (Exámenes Generales Para el Egreso del Tecnico Superior Universitario)	2000	Coordinacion General de Universidades Tecnologicas (CGUT)	CENEVAL	CENEVALによって高等教育機関のキャンパスで実施	所定の日時に実施される全国テストに、高等教育機関が申し込む	一時点	学生個人の結果は非公開。高等教育機関は機関用結果概要を受け取る。学生は自己の得点を示す報告書を受け取る。全国平均以上を得点した者を表彰。
英国	DLHE (Destinations of Leavers from Higher Education, First Destination Supplement に代わる)	2002	連邦政府 : 教育技能省DfEs (Department for Education and Skills) およびその他の政府機関	HESA (Higher Education Statistics Agency)	HESAが質問紙を供給し、各高等教育機関のキャリアセンターが実施	年1度	一時点	集計結果はインターネットで公開される。高等教育機関は機関用結果概要を受け取る。

米国	CAAP (Collegiate Assessment of Academic Proficiency) 1988	該当しない	ACT	高等教育機関が資料を注文して、学生に実施する	柔軟	各要素に複数の形態が準備されており、クロスセクショナルまたは長期的テストが可能	機関用結果概要、学生得点一覧、学生成績表、全国平均以上を得点した学生を表彰、希望に応じて3つまでグループ別集計（たとえば、性別・民族別・専攻別）が可能。テストの要素ごとにサブスコアが算出される。
	MAPP (Measure of Academic Proficiency and Progress, ETS Academic Profileに代わる) 2006	該当しない	ETS, Educational Testing Service, The College Board	高等教育機関が資料を注文して、学生に実施する	柔軟	各要素について複数の形態が準備されており、クロスセクショナルまたは長期的テストが可能	学生得点一覧、機関用結果概要。テストの各要素についてサブスコアが算出される。学年別、カーネギー分類（機関タイプ）別の全国データを公開。Academic Profileの得点に完全に対応している。
	Tasks in Critical Thinking (廃止) 1992	ニュージャージー州高等教育局 (New Jersey Department of Higher Education)	ETS	高等教育機関が資料を注文して、学生に実施する	柔軟	一時点	3つの技能領域のそれぞれについて、熟達水準を上回っている学生の比率を報告。全国データは非公開。
	Major Field Tests (GRE Subject Test) 1990	該当しない	ETS	高等教育機関が資料を注文して、学生に実施する	柔軟	一時点	個人の熟達度スコア、学部平均からの分散を示す学部概要、属性情報を含む機関用結果概要。卒業年度生の受験者の集計表が毎年出版される。学部はテストで提供可能なサブスコアやグループ別集計を購入することができる。
	CLA (Collegiate Learning Assessment) 2002	主要な資金提供者：Carnegie Corporation of New York, William and Flora Hewlett, Ford Foundation, Lumina Foundation, Christian A. Johnson Endeavor Foundation, Teagle Foundation	CAE (Council for Aid to Education)	学生はキャンパスに設置されたインターネット利用可能な場所でオンライン受験をする	年1度	ほとんどの高等教育機関がクロスセクショナルなテストを選択。長期的テストも可能であり、推奨されている。	学生はインターネットで得点報告書を受け取る。高等教育機関は、学生が承諾すれば、個人報告書のコピーを入手することもできる。参加した高等教育機関は機関用結果概要を受け取る。「(SAT得点が)類似した」学生、(カーネギー分類で)同じ種類の機関、米国高等教育システム全体と比較した集計結果が提供される。比較可能な全国データ(大学名は非公開)をオンラインで提供。



米 国 と カ ナ ダ	NSSE (National Survey of Student Engagement) 米2000 (加2004)	The Pew Charitable Trust	Peter Ewellが座長をつとめる全国専門家チーム (NCHEMS)	NSSE (インディアナ大学とNCHEMSの共同出資によるベンチャー企業) によって調査票が配送される	年1度	クロスセクショナル (新入生と卒業年度生)	参加した高等教育機関は機関用結果概要を受け取る。学年別、カーネギー分類 (機関タイプ) 別の全国比較の結果を公表。
	GRE (Graduate Record Examination, General Test) 1966	Graduate Record Examinations (GRE) 理事会 (Association of Graduate Schools, Council of Graduate Schoolsと連携した自律的な理事会)	ETS	アメリカ・カナダ・諸外国各地のテスト会場 (筆記式・コンピュータ式)	柔軟	1時点	学生と受験校 (4校まで) に結果が報告される。

たとえ外部組織によって設計されたアセスメント・ツールが、高等教育機関にとって費用対効果が比較的高いものであっても、学生や高等教育機関が外部テストをキャンパスで実施するという考え方自体に反対する国もあるという。テスト問題が一部の課程にとってレリバンスに欠けていたり、結果が各高等教育機関における教育と学習の改善に直接資するものではないとみなされていたりするからである。高等教育機関における学習とその改善にむけてもっとも役立つのは、属性別の分析であるにもかかわらず、結果がそのような形で提供されていない場合もある (Borden and Owens, 2001)。もっとも、近年では、各高等教育機関が作成した設問を追加したり、各高等教育機関のニーズに即した形でデータを分析したりといった、特別注文のオプションを提供するテスト業者も出現している。

### 3.2.2 アセスメントの頻度

表2に示すとおり、アセスメント・ツールのなかには一時点における学習成果をスナップショットのように示すよう設計されたものもあれば、学生の進歩を長期にわたって測定するよう設計されたものもある。

**一時点テスト (single testing)** では、ある学習期間の終了時に一度だけアセスメントがおこなわれる。EGEL (メキシコ) や MFT (アメリカ) といった、専門分野別の知識に関する成果に焦点をあてた一時点テストであれば、その成果は高等教育機関の課程をとおして教授された内容とある程度関連しているといえよう。ところが一時点テストがより一般的な知識や技能に焦点をあてる場合、その成果が高等教育機関における経験に起因することを証明するのは困難といえよう。一時点テストの結果が、高等教育のインパクトというよりも、学生の累積的な学習成果を反映していると考えられるからである。課程や機関の実績を評価しようとしても、成果が高等教育の経験によると立証することは難しいといわなければならない。

高等教育機関での経験以外にも、極めて多くの要因が学生の学習成果に影響をおよぼしうることが、多くの先行研究より明らかにされている (Pacarella and Terenzini, 2005)。そのため、高等教育機関が学生の学習にもたらす付加価値を測定することをめざすアセスメント・ツールも開発されている。そこではクロス・セクショナル (cross-sectional) ないし長期的 (longitudinal) アセスメントの方法がとられている。

**クロス・セクショナル・アセスメント** : GSA (オーストラリア) や ENADE (ブラジル) など、既にとりあげたアセスメント・ツールのなかには、新入生と卒業年度生をテストの対象としているものがある。このように入学段階の学生と卒業段階の学生を同時にテストする方法は、クロス・セクショナル・アセスメントと呼ばれている (Terenzini, 1989)。

この方法では、新入生は卒業年度生とほぼ同等の入学資格をもつ統制集団とみなされている。そして両群間に観測される差異が、高等教育機関での経験の学生の学習プロセスにもたらす変化として解釈される。

クロス・セクショナル・アセスメントが、同じ学生を繰り返しテストするパネル調査とは異なる点に留意する必要がある。このアプローチの限界の一つは、現在の新入生の主な特性が、現在の卒業年度生の高等教育機関への入学時と同等であるという前提にたっていることである。またこのアプローチでは、継続履修や中退といった、学生による自発的選択の影響が考慮されていない。入学基準の変化が、新入生と卒業年度生の特性に重要な差異をもたらすこともあろう (Terenzini, 1989)。もっとも、パスカレラ (Pascarella, 1987) によると、集団の非同質性の問題は、年齢と入学時の学業適性を統制することによって軽減することができる。

**長期的アセスメント**：同一集団の学習成果を繰り返し測定するアセスメントもある。質問紙調査と比べて、直接的アセスメントでは長期的デザインが用いられることは少ない。もっともアメリカで提供されているテストのなかには、長期的デザインの採用が奨励されているものもある (たとえば CAAP、MAPP、CLA)。これらのテストでは、二時点での実施が可能となるよう、多様な形態が準備されている。理想的な実施形態は、ある学生集団の能力を入学時点で測定し、一定の期間を経た後に、その集団を再度テストするものである。そのように設計された長期的アセスメントの最大の長所は、学生の入学時の特性を考慮することによって、インプット要因の統制が可能になることである。なお入学段階と卒業段階の学生の能力を比較するもう一つの方法として、卒業年度生のテスト結果を、入学試験の成績と比較する方法があげられる。

ただし長期的アセスメントでも、学外での経験など、学習成果に影響をおよぼしうる環境要因が統制されることは少ない。この問題は理論上、新入生と同等の人格的・学業的特徴をもつ高卒者で、高等教育機関に進学しなかった統制集団を同等の期間追跡することで克服することができよう (Terenzini, 1989)。しかしながら、類似した特徴もちつつ、高等教育機関に進学しない生徒のサンプルを確保することは極めて困難である。YITS (カナダ) では、18~20 歳コーホートの発達を、高等教育機関への進学の有無にかかわらず、2年ごとに測定している。そのようなデータを用いて、中等後教育への進学者と非進学者の発達を比較することもできよう。

長期的アセスメントの短所は、調査対象の学生の一部が時間とともに研究から欠落してしまうことである。アセスメントに何年も要し、多大な直接的・間接的な経費がかかることもあげられよう (Terenzini, 1989)。さらにトレンジーニ (1989) は、入学時と卒業時の得点の差異を測定する限界として、「天井効果」に言及している。すなわち入

学時の得点が高かった学生は、低かった学生に比べて、向上の余地が少ない結果、伸び率も小さくなる可能性がある。フェリネとダンテ (Verhine and Dantes, 2005) によると、付加価値アセスメントの問題として、高等教育機関が付加価値スコアを上げるために、新生に低い得点を取るよう働きかけるおそれがある点もあげられる。あるいは高等教育機関は、付加価値スコアを脅かすような劣等生の高学年への進級を認めなくなるかもしれない。

### 3.3 アセスメントのデザイン

このセクションでは、アセスメント・ツールの種類、形式、項目数、アセスメント期間といったアセスメントのテクニカルな側面について整理する。

#### 3.3.1 アセスメント・ツールの種類

表3では、知識と技能の直接的アセスメントと、学生による自己報告などの間接的アセスメントを区別している。この2つのアセスメント・ツールは、さまざまな点で特性を異にしているが、とりわけ「学習という概念からの相対的距離 (relative distance from the construct of learning) (Ewell, 2005, p. 21) の点で大きく異なっている。いずれのアセスメント・ツールを選択するかは、アセスメントをとおしてどのような結果をえようとするのかに規定される。

知識と技能の直接的アセスメントは、学習成果を大規模に測定する際に、もっとも分かりやすいアセスメント・ツールである。比較可能な情報を容易に導くことができるし、アセスメントの条件を細かく統制することで、結果や集団間の比較を客観性の高いものにもすることもできる。(Ewell, 1991)。EGEL (メキシコ)、ENADE (ブラジル)、アメリカで実施されている種々のテストをはじめとする直接的アセスメントは、伝統的に認知的な学習成果に焦点をあててきた。興味深いことに、オーストラリアのGSAには、「対人理解」に関する非認知的要素も含まれている。この要素では、役割、関係性、行動、感情、態度、意欲の機微をとらえる能力や、複雑な対人関係上の問題に対していかなる対応が適切であるかを認識する能力などが測定されている。

本研究で検討する標準化された直接的アセスメントは、通常、日常的な教科指導とは無関係に実施されている。しかし、標準アセスメントの一種で、通常のカリキュラムに埋め込むことのできるものもある。アメリカではMFTなどの標準テストが、総括的な科目に組み込まれることがある。それらの科目で学生は、そこに至るまでのさまざまな科目で教わってきた累積的で統合的な学習成果を証明するよう求められる (Maeroff, 2006)。標準テストを総括的科目に埋め込むことの利点は、テストが履修要

件として自動的に実施されるため、学生をとくに動機づける必要がない点にある。さらに、テストが通常のカリキュラムに埋め込まれていれば、テスト内容と実際に教授された内容との対応関係は高まるであろう (Ewell, 2005)。一方、標準テストに教育内容を対応させることで、テスト準備に過度に焦点をあてたプログラムが出現し、学習環境の多様性が損われる恐れがあると懸念する声もある (Maeroff, 2006)。

自己アセスメントは、学習そのものではなく、学習の間接的な証拠に着目する。意識調査では、学生や卒業生に、自己の学習成果を評価させたり、中・長期的な行動面の成果を報告させたりしている。たとえば NSSE (アメリカ) では、学生がさまざまな学習活動にいかに関与したかが測定されている。CEQ (オーストラリア) では、履修科目が知識と技能、対人関係理解を高めるうえでいかに役立ったかを、学生に評価させている。

卒業生を対象とした意識調査も、就職率や進学といった学習の二次的指標に関する情報を収集する手段となりうる。NGS (カナダ) や DLHE (イギリス) などの意識調査において学生からえられた情報は、カナダ統計局やイギリスの高等教育統計局といった政府・教育機関が所有する既存の統計記録にリンクさせることができる。そうした合体データを用いて、教育と雇用の関係に関する有益な情報をえることができる。

### 3.3.2 形式

表3に示す直接的アセスメントの形式は、選択式と記述式に区別される。直接的アセスメントのほぼすべてで、両形式が組み合わせて採用されている。

**選択式アセスメント**では、項目ごとに明確な回答様式が備わっており、自動的に数値スコアが割り当てられている (Ewell, 2005)。GSA (オーストラリア)、EXANI-III (メキシコ)、EGEL (メキシコ)、EGETSU (メキシコ)、CAAP (アメリカ)、MAPP (アメリカ)、MFT (アメリカ) をはじめとする多くのアセスメントで採用されている標準多肢選択式テストが、その好例である。選択式テストにおけるアセスメントの設計者の役割は、テスト項目と採点基準を作成することに留まり、学生の答案の採点に携わることはない。このようなアセスメントは、比較可能性の高い結果をもたらすが、その欠点として、狭い範囲の学習成果しかテストできない点があげられる。

**記述式アセスメント**では、自動的に採点することが必ずしも可能ではない、記述式の解答を作成したり、独自の解答を述べたりすることが求められる。そうしたアセスメントでは、解答の質の判定する専門家に重要な役割が委ねられている。アセスメントには大抵、採点指針や水準別の解答の見本が準備されているが、人間の判断が伴う標準的な

バラつきを回避することはできない (Ewell, 2005)。一方、このようなアセスメント方法は、文章表現力、総括力、組織力といった広い範囲の学習成果をとらえることができる利点をもつ。GSA (オーストラリア)、ENADE (ブラジル)、CAAP (アメリカ)、MAPP (アメリカ) で出題される記述問題や小論文などが、その例としてあげられる。

### 3.3.3 アセスメントにおける質の基準

**目標に準拠したアセスメント** (criterion-referenced assessment) : 達成度を絶対的基準にもとづいて評価する方法は、**目標に準拠したアセスメント**と呼ばれている。「質」または「習熟度」について、絶対的基準やベンチマークを採用しているアセスメントは多くない。メキシコの2種類の卒業試験 (EGEL と EGETSU) は、あらかじめ設定された、特定のテスト得点に対応する、望ましい達成度の基準に照ら合わせて習熟度が測定されている数少ない例である。大学の学位を必要とする職業が要請するコンピテンスの複雑さに鑑みて、多くの専門家が、高等教育におけるテストにそうした基準を設けることの難しさを強調している (Verhine and Dantas, 2005)。

EGEL や EGETSU のような目標に準拠したアセスメントの実践においては、卓越性の基準ではなく、達成度の「最低基準」が設定されている場合が多い。最低基準は、専門的な資格・免許試験でしばしば用いられており、特定の専門分野におけるコンピテンスを証明する「最低限」の知識と技能を定義している。たとえば医学生を対象とする EGEL (メキシコ) では、新規に卒業した「すべての若い医師」が習得していなければならない知識と技能をテストしている (Hughet, 2000)。そのねらいは、卒業生の能力が最低レベルに達しない高等教育機関の存続を阻むことにある (Hughet, 2000)。

**集団に準拠したアセスメント** (norm-referenced assessment) : 高等教育における標準アセスメントのほとんどは、受験者間の比較にもとづいて受験者を分類している (Verhine and Dantas, 2005)。このようなアセスメントは**集団に準拠したアセスメント**と呼ばれている。集団に準拠したテストの結果は、他の学生との比較、または同じ学生の異なる時点との比較においてのみ、意味をもつ (ETS, 2006)。それらは高等教育機関の卒業生の絶対的基準に照らし合わせた習熟度ではなく、相対的な位置づけを示すにすぎない (Pascarella and Terenzini, 2005)。マクミラン (McMillan, 1988) によると、そのような相対的な習熟度をどのように評価するか判断することは難しく、知識と技能が相対的に高まっていても、達成度の絶対的基準を満たしていない場合もありうる。

表3. アセスメントのデザイン

国	テスト名	開始年	アセスメント	形式	項目数	アセスメントの期間
オーストラリア	GSA (Graduate Skills Assessment)	2000	直接的	多肢選択式・記述式	多肢選択式1問、記述式2問(説明・論述)	3時間(選択式2時間・記述式1時間)
	CEQ (Course Experience Questionnaire, Graduate Destination Surveyの一部)	1993	間接的	質問紙調査	25問(履修科目の満足度、一般的技能の習得に関する評価)	5~10分
	GDS (Graduate Destination Survey)	1972	間接的	質問紙調査	就業について19問、進学について6問	5~10分
ブラジル	ENC/Prova (Exame Nacional de Cursos)	1995-2003	直接的	詳細不明	詳細不明	4時間
	ENADE (Exame Nacional de Desempenho dos Estudantes)	2004	直接的	客観問題・記述式	一般問題10問、専門分野別問題30問	4時間
カナダ	NGS (National Graduate Survey, Follow-up Survey of Graduates)	1978	間接的	電話インタビュー(CATI)	18セクション、113頁(該当部分のみ回答)	詳細不明
	YITS (Youth in Transition Survey)	2000	間接的	電話インタビュー(CATI)	詳細不明	詳細不明
メキシコ	EXANI-III (Examen Nacional de Ingreso al Posgrado)	1997	直接的	多肢選択式	120問(66問54%は一般的知的能力、54問46%は情報の活用能力を問う)	4時間
	EGEL (Examen General Para el Egreso de la Licenciatura)	1994	直接的	多肢選択式	100~300問(専門分野によって異なる)	8~10時間(週末に複数セッション)
	EGETSU (Exámenes Generales Para el Egreso del Técnico Superior Universitario)	2000	直接的	多肢選択式	250問(87問35%は一般的要素、163問65%は専門分野の要素)	6時間(3時間セッション2回)

英国	DLHE (Destinations of Leavers from Higher Education, First Destination Supplementに代わる)	2002	間接的	質問紙調査	56問	約15分
米国	CAAP (Collegiate Assessment of Academic Proficiency)	1988	直接的	多肢選択式・記述式	6種類のなかから選択 (各72問以内)、9問まで機関独自の質問を追加可能	各セクション40分
	MAPP (Measure of Academic Proficiency and Progress, ETS Academic Profileに代わる)	2006	直接的	多肢選択式・記述式	長編108の多肢選択式問題 (100分)、短編36問。小論文問題も実施可能	2種類の標準形式 (各2時間)、6種類の短縮形式 (各40分)
	Tasks in Critical Thinking (廃止)	1992	直接的	記述式・実践的問題解決課題	詳細不明	各問題90分
	Major Field Tests (GRE Subject Test)	1990	直接的	多肢選択式	詳細不明	2時間 (MBAは3時間)
	CLA (Collegiate Learning Assessment)	2002	直接的	実践的作業	詳細不明	90分
米国とカナダ	NSSE (National Survey of Student Engagement)	米2000加2004	間接的	質問紙調査	取り組みに関する5つのセクション (90問)	約15分
	GRE (Graduate Record Examination, General Test)	1966	直接的	多肢選択式・記述式	詳細不明	3時間 (コンピュータ式)



### 3.4 対象者

このセクションでは、アセスメントのターゲット人口、サンプルの抽出法、テストの適用範囲、学生や高等教育機関への参加の動機づけについて情報を提供することで、アセスメント対象者の特徴を整理する。

#### 3.4.1 ターゲット人口

表4に示すとおり、本稿で検討するアセスメント・ツールのほとんどが、学習期間の終わりに学生の成果を測定する仕組みをとっている。課程や機関レベルの直接的アセスメントでもっとも頻繁にターゲットとされているのは、学士課程の修了を控えた卒業年度生である。MAPP（アメリカ）、GSA（オーストラリア）、ENADE（ブラジル）などがその例としてあげられる。

学士課程段階で学習成果を測定する利点として、この段階では多くの課程が、一般的学業技能の育成をはじめとした目標を共有している点をあげることができる。大学院段階の学習成果アセスメントは、極めて特殊で多様なアセスメント・ツールを必要とするため、設計するのがより困難である。さらに大学院と比較して、学士課程の在籍者数は多いため、学士課程における学習のアセスメントには、より大きな社会的関心がよせられている。

卒業後の個人を対象とした課程や機関レベルのアセスメントは、ほとんど存在しない。それはおそらく、大学を離れてから大規模テストを実施することも、卒業生にテストに参加するよう動機づけることも難しいためであろう。卒業後の学生に対して直接的アセスメントがおこなわれる場合、通常それは個人レベルの結果に注目するものであり、大学院入学試験や資格試験など、受験者にとって重大な結果をもたらすものである。メキシコのEGELは、単位をすべて取得済みであることを条件に、卒業の有無を問わず、受験することができる。卒業証書に加えて全国的に認知された資格を取得できる可能性が、受験にむけて卒業生を動機づける要因となっているのであろう。

履修科目やキャンパスでの経験に関わる意識調査は、学生と卒業生のいずれをも対象に実施されうる。NSSE（カナダとアメリカ）は新入生と卒業年度生の両方に配布されており、オーストラリアのCEQは卒業後4ヶ月目に卒業生に送付される。雇用や進学といった学生による成果の二次的指標に着目する意識調査は、当然ながら卒業後しばらく経った個人をターゲットとして実施されている。GDS（オーストラリア）、NGS（カナダ）、DLHE（イギリス）などがその例に含まれる。

表4. アセスメントの対象者

国	テスト名	開始年	対象者の特徴			動機づけ	
			ターゲット人口	サンプル抽出法	テストの適用範囲	学生	機関
オーストラリア	GSA (Graduate Skills Assessment)	2000	新入生・卒業年度生	限定的・希望者	毎年、多様な専門分野から約20大学、学生約2,000人が参加	就職する際の履歴書にGSA得点を記載できる	入学段階：結果は学力不振の学生を識別し、支援を提供するために使用することができる。卒業段階：結果は卒業・大学院入学の補足的な基準として活用することができる。機関によってもたらされた「付加価値」、各高等教育機関・課程における全般的な学習の質に関する情報として解釈することができる。
	CEQ (Course Experience Questionnaire, Graduate Destination Surveyの一部)	1993	前年度に何らかの高等教育資格を修了した卒業生全体	ターゲット人口全体	オーストラリアの高等教育機関の卒業生（留学生を含む）全員に調査票が送付される。回収率は60～65%	—	高等教育機関は卒業生の結果について報告をうける。学内の目的に即して結果を分析・報告することができる。
	GDS (Graduate Destination Survey)	1972	前年度に何らかの高等教育資格を修了した卒業生全体	ターゲット人口全体	オーストラリアの高等教育機関の卒業生（留学生を含む）全員に調査票が送付される。回収率は60～65%	—	高等教育機関は卒業生の結果について報告をうける。学内の目的に即して結果を分析・報告することができる。
ブラジル	ENC/Prova (Exame Nacional de Cursos)	1995-2003	卒業年度生	ターゲット人口全体	事前に指定された専門分野の卒業年度生全員。2003年には6,500課程に在籍する学生46万人（卒業年度生全体の70%）が受験	参加を学士号授与の要件として義務化。成績優秀者は表彰	参加は義務。課程・機関は、他機関の同じ専門分野と比較する形で、学生の成績に関する報告をうける。
	ENADE (Exame Nacional de Desempenho dos Estudantes)	2004	事前に指定された専門分野における初年次生（課程の7～22%を終了）と卒業年度生（課程の80%以上を終了）	有資格者のリストから無作為抽出（20人以下の課程は全数）	2004年の無作為抽出サンプルには、2,184課程の学生140,340人が含まれていた（初年次生の51%、卒業年度生の69%）。無作為抽出で選ばれなかった2,839人の希望者が任意で参加	無作為抽出された学生にとって参加は義務（学士号授与の要件）。成績優秀者は表彰	参加は義務。課程・機関は、学生の成績について絶対的・相対的評価の形で報告をうける。Provaと異なり、ENADEでは結果はメディアに公表ず、インターネットをとおして個別に報告される。

カナダ NGS (National Graduate Survey, Follow-up Survey of Graduates)	1978	すべての公立の高等教育機関の卒業生を対象に卒業後2年目を実施。一部には5年目に追跡調査を実施。	地方・課程・専門分野ごとに正確な予測値がえられるよう、層化抽出法を採用	初期サンプル35,000～40,000人。参加任意。平均回収率は65%。	— アセスメントでは個々の高等教育機関に焦点が当てられるのではないため、高等教育機関にとって負のインパクトはない。高等教育機関は、卒業生の就職状況について報告をうける。学内の目的に即して結果を分析・報告することができる。
YITS (Youth in Transition Survey)	2000	15歳、18～22歳人口全体（学生・卒業生に限定されない）	カナダ労働力調査サンプルに対して多段階層化抽出法を採用	参加は任意。初期サンプルは29,000人。18～20歳コーホートの回収率は約80%	— 該当しない
EXANI-III (Examen Nacional de Ingreso al Posgrado)	1997	大学院進学予定者	このテストを採用している大学院への入学を希望する学生全員	このテストを採用している大学院への入学・奨学金受給学を希望する学生全員。2006年には13,604人が受験	一部の大学院で参加が義務化されている。奨学金に応募する際に、自己の能力を証明する手段となる アセスメントでは個々の高等教育機関に焦点が当てられるのではないため、高等教育機関にとっての負のインパクトはない。大学院は学生の成績について、比較可能な情報をえることができる。
メキシコ EGEL (Examen General Para el Egreso de la Licenciatura)	1994	非工学系の学生で、課程をすべて修了した者（卒業の有無にかかわらず）	高等教育機関または学生が任意で参加。高等教育機関が参加を申し込んだ場合、高等教育機関がサンプルの数と生徒の特徴を決定する	詳細不明	全国的に通用する補足的な資格をえることができる。全国平均を上回る学生を表彰 高等教育機関は卒業生の結果について報告をうける。学内の目的に即して結果を分析・報告することができる。結果を卒業や進学の補足的な基準として利用することもできる。
EGETSU (Exámenes Generales Para el Egreso del Técnico Superior Universitario)	2000	工学系の学生で、課程をすべて修了した者（卒業の有無にかかわらず）	候補者全員	全国48の工学系高等教育機関の卒業年度生全員	全国的に通用する補足的な資格をえることができる。全国平均を上回る学生を表彰 高等教育機関は卒業生の結果について報告をうける。学内の目的に即して結果を分析・報告することができる。結果を卒業や進学の補足的な基準として利用することもできる。

<p>英国</p> <p>DLHE (Destinations of Leavers from Higher Education, First Destination Supplementに代わる)</p>	<p>2002</p> <p>公的資金をうけた高等教育機関でフルタイム・パートタイムで就学し、学位・資格を取得した卒業生 (卒業後約6ヵ月)</p>	<p>候補者全員</p>	<p>候補者全員に調査票が送付される。2002年度にはフルタイム有資格者の77% (251,300人) およびパートタイム有資格者の70% (60,900人) が回答</p>	<p>—</p>	<p>アセスメントでは個々の高等教育機関に焦点が当てられるのではないため、高等教育機関にとって負のインパクトはない。高等教育機関は、卒業生の就職状況について報告をうける。学内の目的に即して結果を分析・報告することができる。</p>
<p>米国</p> <p>CAAP (Collegiate Assessment of Academic Proficiency)</p>	<p>1988</p> <p>あらゆるタイプの学生</p>	<p>州または高等教育機関がアセスメントを実施するかを決定し、学生サンプルの規模と特徴を規定する</p>	<p>1988～2001年の間、600機関45万人が参加</p>	<p>高等教育機関が動機づけを提供している場合もある。テストの動機づけに関する項目より、学生がどれだけ真剣にテストを受験したかを判断することができる</p>	<p>高等教育機関は卒業生の結果について報告をうける。学内の目的に即して結果を分析・報告することができる。結果は機関によってもたらされた「付加価値」、および各高等教育機関・課程における全般的な学習の質に関する情報として解釈することができる。</p>
<p>MAPP (Measure of Academic Proficiency and Progress, ETS Academic Profileに代わる)</p>	<p>2006</p> <p>あらゆるタイプの学生</p>	<p>州または高等教育機関がアセスメントを実施するかを決定し、学生サンプルの規模と特徴を規定する</p>	<p>MAPPのもととなるAcademic Profile (1992～2006年) は、375の機関と100万人の学生によって利用されてきた</p>	<p>高等教育機関が動機づけを提供している場合もある</p>	<p>高等教育機関は卒業生の結果について報告をうける。学内の目的に即して結果を分析・報告することができる。結果は機関によってもたらされた「付加価値」に関する情報として解釈することができる。一般教育における学生の成績を、他の課程・機関との比較のうえで明らかにする。</p>

米 国	Tasks in Critical Thinking (廃止)	1992	あらゆるタイプの学生	州または高等教育機関がアセスメントを実施するかを決定し、学生サンプルの規模と特徴を規定する	1992～2001年にかけて、35機関（各200～500人の学生）で実施された。	高等教育機関が動機づけを提供している場合もある	高等教育機関は卒業生の結果について報告をうける。学内の目的に即して結果を分析・報告することができる。結果は機関によってもたらされた「付加価値」に関する情報として解釈することができる。批判的思考における学生の成績を、他の課程・機関との比較のうえで明らかにする。
	Major Field Tests (GRE Subject Test)	1990	卒業年度生（4年制大学）	州または高等教育機関がアセスメントを実施するかを決定し、学生サンプルの規模と特徴を規定する	毎年500大学余りが利用。1999年度には、606機関の1,000以上の学部が約7万人の学生に対して実施	総括的科目として、あるいは最終学期に卒業要件としてテストが実施される場合が多い	高等教育機関はアセスメントをシラバスに組み込み、テストを卒業要件とすることができる。学内の目的に即して結果を分析・報告することができる。専門分野における学生の成績を、他の機関との比較のうえで明らかにする。
	CLA (Collegiate Learning Assessment)	2002	新入生・卒業年度生	毎年各機関より新入生100人と卒業年度生100人を抽出。試験問題全体を個別の要素に分割し、無作為抽出された学生のサブサンプルに適用することで（マトリックス・サンプリング）、テスト時間を最小限に抑制しながら、多様なツールと内容を完全に網羅することができる	2002～2005年に134大学が参加	高等教育機関のなかには参加者に贈り物・賞・記念品を与えている場合もある	高等教育機関は結果の報告をうける。指導方法、性別、人種・民族性、社会経済的地位別に分析することで、マイノリティ学生の成功を支援することができる。結果は機関によってもたらされた「付加価値」に関する情報として解釈することができる。CLAデータベースで大学名は明らかにされない。

米国とカナダ	NSSE (National Survey of Student Engagement)	米2000 加2004	新入生・卒業年度生 (4年制大学)	新入生及び卒業年度生全員のリストから無作為抽出	2000年以来、北米の1,000校余りの大学が学生116万人に対してNSSEを実施。サンプル・サイズは学生数によって規定される。(各機関450~1,000人)	—	高等教育機関は学生の結果について報告をうける。学内の目的に即して結果を分析・報告することができる。学士課程の経験のなかで、大学改革によって改善可能な側面を明らかにすることができる。
	GRE (Graduate Record Examination, General Test)	1966	大学院進学予定者	このテストを入学要件とする大学院進学希望者全員	詳細不明	アメリカおよび英語圏の国々における多くの大学院において、入学要件となっている	アセスメントでは個々の高等教育機関に焦点があてられるのではないため、高等教育機関にとって負のインパクトはない。大学院レベルの高等教育機関は、学生の成績について比較可能な情報をえることができる。

### 3.4.2 サンプルの抽出

直接的アセスメントをもちいてターゲット人口の悉皆調査を行うには、アセスメントが義務化されているか、個人にとって重大な結果をもたらす場合にのみ可能である。ブラジルの ENC-Provao は、学生人口全体を対象とした直接的アセスメントの唯一の事例であったが、2003 年に打ち切られた。そのような大規模な集団に対するテストの実施は、非常に多くの人的・経済的資源を必要とする。悉皆調査は、実施コストが比較的安価な意識調査などのような間接的アセスメントにより適しているといえよう。さまざまな卒業生向けの意識調査（カナダの NGS、イギリスの DLHE）で、ターゲット人口全体に質問紙が送付されている。

抽出調査は、ターゲット人口全体を代表するサンプルにテストを実施することで、はるかに低いコストで学生の達成度に関する正確な予測値をえるアセスメント方法である。正確な結果を導くためにサンプルは十分な規模を満たし、学生人口の多様な集団をその比率に応じて代表していなければならない。抽出調査は明らかに、悉皆調査よりも安価である。しかしながら一部のステークホルダーにとって、信頼性に欠けると受け止められる場合もある（Terenzini, 1989）。抽出調査のアプローチは、ブラジルの一部のステークホルダーによって批判されたが、その理由は高等教育機関がテストで高得点をとれそうな学生だけを採用した場合、結果がゆがめられてしまうというものであった（Verhine and Dantes, 2005）。

### 3.4.3 動機づけ

プログラムの質または教育機関の達成度の測定をめざすアセスメントに共通する問題は、学生に最善を尽くさせるための動機づけに欠くことである。ほとんどのアセスメントにおいて、学生は成果をめぐる直接利害を経験するわけではない。自らの高等教育機関の評価が低ければ、学生にも最終的には負の影響がおよぶものの、学生にとってテスト得点は直接の関心事ではない。

学生にテストを受験するよう動機づけるためのさまざまな試みが展開されてきた。Provao（ブラジル）では、受験が卒業要件に含まれていた。ただし興味深いことに、義務化されていたのは、受験会場に来ることであり、問題に解答することではなかった（Verhine and Dantes, 2005）。その結果として多くの学生が、解答用紙を白紙で提出したり、真剣に取り組まなかったりした（Verhine and Dantes, 2005）。イギリスで現在パイロット調査として実施されている、新たな長期的研究 DLHE では、受験者一人につき 1 ポンドを慈善事業に寄付する約束で、学生の受験の動機づけを図っている（HESA ホームページ、2006）。CLA を採用しているいくつかの高等教育機関からは、学生の参加

を促すために、少額の謝礼金や贈り物を提供していることが報告されている。

こうしたアセスメントで学生がよい成績をあげるよう奨励する一つの方法として、最高点に賞や証書など、学生にとって利益となるものを用意する方法があげられる。ENADE（ブラジル）では、テストで優秀な成績を修めた学生に、賞と奨学金を授与している。オーストラリアとメキシコの政府機関は標準テスト（オーストラリアのGSA、メキシコのEGEL/EGETSU）を、学生が就職や進学の際に履歴書に記載できる、全国的に認知された特別資格として設計している。オーストラリアとメキシコの政府は、雇用者や大学院が学生の能力を判断する資料として、テスト結果の提示を求めることを奨励している。

しかしながら、課程や機関レベルのアセスメントに、学生にとって重大な結果をもたらす要素を含めた場合、動機づけによる偏りが生じる可能性もある。テストが将来にとって重要であるならば、学生はテストのためにとくに準備をするであろう。その場合テスト結果は、高等教育機関における経験が学習をいかに促したかではなく、個々の学生が特定のテストのためにいかによく準備したかを反映することになる。

### 3.5 アセスメントの目的と結果

このセクションでは、アセスメントの目的に注目し、個々の学生、課程や機関、または教育制度のいずれを測定しようとしているのかを明らかにする。いかなる結果が追究されており、それらが各ステークホルダーによってどのように利用されるのかについて整理する。

#### 3.5.1 アセスメントの焦点

いずれの学習成果アセスメントでも、個々の学生の結果が一次データとして使用されている。しかしながら、アセスメントの目的によっては、学生がアセスメントの主たる焦点ではない場合もある（Terenzini, 1989）。テストは、個人レベルのアセスメントを目的とするものと、集団レベルのアセスメントを目的とするものに区別することができる。アセスメント・ツールの多くで、複数のレベルが組み合わせられて活用されている。

**個人レベルのアセスメント**では、個々の学生の成績に主な焦点があてられ、集団レベルの成績はあまり重視されない。たとえば GRE（アメリカ）や EXANI-III（メキシコ）などの大学院入学試験では、学生選抜に関わる決定に資するために、個々の学生の成績が検討される。個人レベルのアセスメントは、成績の悪い学生を識別し、追跡して、特別な支援を提供するために実施されている場合もある。たとえばオーストラリアのGSAは、集団レベルの結果を提供する一方で、個人レベルの要素も有している。



大学院入学試験のような個人レベルのアセスメントを用いて、学生が履修してきた学士課程の質に関する結論を導くことが提案されている。しかしながら、個人レベルのアセスメントは通常、教育のすべての過程をとおして習得されてきた学習の累積的な成果を測るものである。特定のプログラムや教育機関が、学習成果にいかに関与したかには、あまり注目しないものである。

**集団レベルのアセスメント**では、集計結果に焦点があてられる。個々の学生の結果そのものをアセスメントの目的とみなすのではなく、平均値としての集団レベルの指標を作成するデータとして使用している。ここで「集団」とは、課程、機関、教育制度といった、学生の学習成果を集計するさまざまな単位を意味する。集団レベルのアセスメントは、課程や機関が学生の学習成果にどれほど影響をおよぼしたのかを明らかにするために利用することができる。

**課程レベルのアセスメント**では、異なる機関で、同じ課程を履修している学生の達成度の差異が検討される。課程レベルのアセスメントでは、専門分野別の知識と技能に焦点があてられることが多い。このようなアセスメントをとおして、異なる教授法を用いて、同じ学習成果の達成をめざしている課程の比較が可能になる。たとえば異なる高等教育機関や課程に在籍する、同じ専門分野の学生の達成度の比較を行うブラジルの ENC-Prova (2003年に終了) や MFT (アメリカ) などの専門分野別のテストが例としてあげられる。

**機関レベルのアセスメント**では、履修しているプログラムに関わらず、在籍している機関別に集計された学生の学習成果に焦点があてられる。プログラム全体に共通する一般的な学業能力がテストされることもあれば、各課程の専門分野別の問題が含まれることもある。たとえばオーストラリアの GSA、ブラジルの ENADE、アメリカの各種テストが例としてあげられる。

**制度レベルのアセスメント**では、学生の達成度における国家間の差異が検討される。社会的要因、構造的特徴、教育政策の枠組みなどの制度レベルの要因が、学生の達成度におよぼす影響に焦点があてられる。本稿で検討するアセスメント・ツールのなかで唯一、複数の国家の高等教育機関で使用されているのは、アメリカおよびカナダの高等教育機関で実施されている NSSE (National Survey of Student Engagement) である。カナダとアメリカにおける学生の取り組みの比較にもとづく制度レベルの差異が、ホームページおよび政府報告書で公表されている。なお GRE (アメリカ) のような、極めて多数の教育制度の出身者が受験する大学院入学試験から、制度レベルの比較を行うことも潜在的に可能といえよう。

### 3.5.2 アセスメント結果の使途

**形成的評価**の目的は、直後の行動を導くうえで、すぐに役立つ結果をえることである (Ewell, 1991)。フォルクヴァイン (Volkwein, 2003, p. 7) によると、形成的評価とは「指導が学習に影響を与え、学習が成果に影響を与え、成果のアセスメントが指導、そして最終的には学習に影響と改善をもたらすフィードバック・ループ」といえる。

個人のレベルでは、形成的評価とは、学生に知識と技能における進歩や成長について定期的にフィードバックすることを意味する (Terenzini, 1989)。さらに教師が個々の学生の発達ニーズを見極め、適切な指導計画によってニーズに適切に対応するために、形成的評価を利用することもできる。GSA (オーストラリア) は、この形成的アセスメントの好例である。すなわち入学時のテストを用いて達成度の低い学生を識別し、特別な支援とフォローアップを提供することがめざされている。

課程や機関レベルでは、形成的評価はプログラムの改善と、指導の有効性の向上に資することができる。高等教育機関レベルの形成的評価は、体系的な自己評価が機関における継続的な学習と改善につながるという前提にもとづいている (Astin, 1977 ; Bowen, 1977 ; Ewell, 1984)。高等教育機関は、形成的評価をおこなうために、市販のアセスメント・ツールを利用することができる。アメリカでは現在そのような状況にあり、CAAP や MAPP などのテストが利用されている。形成的評価のデザインは、機関による自己省察と証拠にもとづく (エビデンス・ベースドな) プログラム・デザインを促す。高等教育機関が達成度のベンチマークを定め、傾向を分析し、個々のプログラムを評価し、カリキュラムを編成・改善することを可能にするのである。

**総括的評価**は、組織的上位の権威に対するアカウンタビリティ要請に応えることを目的とし、判定や記録のために達成度を評価することを意味する (Ewell, 2005)。

個人のレベルでは、総括的評価は有資格者と無資格者を区別するといった選別機能を果たすことがある (Terenzini, 1989)。標準テストは、プログラムのさまざまな時点で、そうした選別の目的で利用される。さらに個人のレベルでは、総括的評価はその個人があるポストに適任であることを、雇用者に対して証明しなければならない際にも使われる。そのような目的で、GSA (オーストラリア)、EGEL および EGETSU (メキシコ) の結果は、学生によって利用されている。

課程や機関のレベルでは、アセスメントの集計結果が、国または州によるより広範なアカウンタビリティの枠組みのなかで使用されることがある。課程や機関にとっての事

態の重大さは、国によって異なる。アクレディテーション協会が、学生の学習成果に着目している国はほとんどない。しかしながら、高等教育機関が適格認定を受け、公的資金を受け取る前提条件として、機関ごとになんらかの成果アセスメントに取り組むことが義務づけられている国も一部ある。機関の実績指標をえる目的だけのために、全国規模または全州規模の標準アセスメントが実施されることは、未だ極めて稀といえる (Ewell, 2005)。

学習成果に関する標準テストの結果を、適格認定や資源配分に直接連動させているのはブラジルと、アメリカの一部の州に限られている。オーストラリアをはじめとする一部の国では、政府が結果から全国学生の達成水準を明らかにし、政策を決定したり高等教育のニーズへの対応を計画したりするための資料としている。

集計されたアセスメントの結果は、進学希望者、その保護者、雇用者への情報提供のために使用されることもある。一部の高等教育機関や政府は、質についての理解を、評判や地位ではなく、効果的な学習や「目標への適合性」に注目するものへとシフトさせようと努めている (Baldwin, 2000)。学習成果アセスメントの結果は、高等教育機関の質について、教育や学習の観点から重要な証拠を提供してくれると考えられる。

表5. アセスメント結果の目的と用途

国	テスト名	開始年	焦点	結果	結果の用途
オーストラリア	GSA (Graduate Skills Assessment)	2000	機関	入学段階・卒業段階の学生の一般教育の成果	高等教育機関：入学段階では学業不振の学生を識別し、支援を提供する。卒業段階では大学院レベルの課程への入学に関する補足的な基準となる。動向のベンチマークと分析、プログラムの効果や改善の長期的な記録、学生の成績を全国標準と比較、カリキュラムの開発・改善、学生による上級科目の履修資格の判定。政府：全国の高等教育機関における学習成果の質に関する情報を収集、大卒者の技能に関する全国的・国際的ベンチマークを開発。雇用者：標準的な採用ツールとして、企業がテストを利用することを政府が奨励している。
	CEQ (Course Experience Questionnaire, Graduate Destination Surveyの一部)	1993	課程・機関	教育と学習に対する卒業生の満足度、高等教育機関での経験に関連する学業技能の習得についての自己報告	高等教育機関：ベンチマークと動向調査、プログラム評価、カリキュラム開発・改善、全国的なアカウンタビリティ・データの提供。政府：高等教育の質保証、学生による選択に資する情報提供、高等教育のニーズ調査と対応。2005年以降、CEQの結果を「学習・教育向上予算(Learning and Teaching Performance Fund)」のパフォーマンスにもとづく配分に活用。
	GDS (Graduate Destination Survey)	1972	課程・機関	就業と進学に関する情報	高等教育機関：ベンチマークと動向調査、プログラム評価、就業・進学実績を最適化するためのカリキュラム開発・改善、全国的なアカウンタビリティ・データの提供。政府：高等教育の質保証、学生による選択に資する情報提供、高等教育のニーズ調査と対応。2005年以降、GDSの結果を「学習・教育向上予算(Learning and Teaching Performance Fund)」のパフォーマンスにもとづく配分に活用。
ブラジル	ENC/Prova (Exame Nacional de Cursos)	1995-2003	課程	他機関の同じ専攻分野の学生の成績の比較	高等教育機関：よい結果は商業目的で利用(広告宣伝)、よい得点を維持して悪い得点を向上させるよう、教員と学生を動機づける要因となっている。政府：2001年よりアクセディテーションの要件として参照されているが、懲罰的な対応はほとんど取られていない。一般社会：高等教育機関における学習の質について入学志願者や社会に情報を提供する目的で、メディアによって利用されてきた。
	ENADE (Exame Nacional de Desempenho dos Estudantes)	2004	課程・機関	入学段階・卒業段階の学生の認知的成果の差異	学生：雇用者に対して、自己の成績を全国標準と照らし合わせて証明。高等教育機関：ベンチマークと動向調査、プログラム評価、カリキュラム開発・改善、全国的なアカウンタビリティ・データの提供。政府：高等教育の評価の一指標。雇用者：専門分野における能力の客観的証拠として就職希望者に開示を求めることができる。

カナダ	NGS (National Graduate Survey, Follow-up Survey of Graduates) 1978	課程・機関	卒業生による労働市場参入に関する情報、教育と労働市場における成果の関連	高等教育機関：ベンチマークと動向調査、プログラム評価、就業・進学実績を最適化するためのカリキュラム開発・改善、全国的なアカウンタビリティ・データの提供。政府：高等教育のニーズ調査と対応。
	YITS (Youth in Transition Survey) 2000	教育システム	青少年による学校（高等教育機関）から労働市場へのトランジションに関する情報	Human Resources Skills Development Canada：青年の教育・訓練・就業を中心とした生活のパターン、影響をおよぼす要因、主要なトランジションについて情報を収集する。政策やプログラム開発に資する。教育者、政策立案者、支援団体に情報を提供する。
メキシコ	EXANI-III (Examen Nacional de Ingreso al Posgrado) 1997	機関タイプ・地域	累積的学業能力、大学院での研究に必要な一般学業技能の習得	高等教育機関：各学生の結果にもとづいて成績を比較検討し、学生の入学選抜や奨学金の給付に関する意思決定を行う。
	EGEL (Examen General Para el Egreso de la Licenciatura) 1994	学生	累積的成果（専門分野別）、専門職に必要な必須知識と技能の習得	学生：雇用者に対して、自己の成績を全国標準と照らし合わせて証明。高等教育機関：ベンチマークと動向調査、プログラム評価、カリキュラム開発・改善、卒業・資格授与の補足的な基準としている高等教育機関もある。雇用者：専門分野における能力の客観的証拠として就職希望者に開示を求めることができる。
	EGETSU (Exámenes Generales Para el Egreso del Tecnico Superior Universitario) 2000	学生・課程・機関	累積的成果（専門分野別）、専門職に必要な必須知識と技能の習得	学生：雇用者に対して、自己の成績を全国標準と照らし合わせて証明。高等教育機関：ベンチマークと動向調査、プログラム評価、カリキュラム開発・改善、卒業・資格授与の補足的な基準としている高等教育機関もある。雇用者：専門分野における能力の客観的証拠として就職希望者に開示を求めることができる。
英国	DLHE (Destinations of Leavers from Higher Education, First Destination Supplementに代わる) 2002	課程・機関	就業と進学に関する情報	高等教育機関：ベンチマークと動向調査、プログラム評価、就業・進学実績を最適化するためのカリキュラム開発・改善、全国的なアカウンタビリティ・データの提供。政府：高等教育のニーズ調査と対応。

	CAAP (Collegiate Assessment of Academic Proficiency)	1988	学生・課程・機関	一時点/付加価値デザインによって、一般学業技能の累積的成果/大学在学中に習得した一般学業成果	高等教育機関：アクレディテーションおよびアカウンタビリティの要件を満たす、ベンチマークと動向調査、プログラムの効果と改善の長期的な記録、学生の成績を全国標準と比較、カリキュラムの開発・改善、学生による上級科目の履修資格の判定。州：CAAPは教育システム全体を一定期間モニターするために使用されており、公共へのアカウンタビリティの枠組みのなかで、進歩のベンチマークとして活用されてきた。州全体の結果は、報告書にまとめられ、他州と比較することができる。
	MAPP (Measure of Academic Proficiency and Progress, ETS Academic Profileに代わる)	2006	学生・課程・機関	大学在学中に習得した一般的学業技能	高等教育機関：アクレディテーションおよびアカウンタビリティの要件を満たす、ベンチマークと動向調査、プログラムの効果と改善の長期的な記録、学生の成績を全国標準と比較、カリキュラムの開発・改善、学生による上級科目の履修資格の判定、学生に対する成績指導。
米国	Tasks in Critical Thinking (廃止)	1992	学生・課程・機関	実践的作業にもとづく大学レベルの高次思考スキルの評価	高等教育機関：アクレディテーションおよびアカウンタビリティの要件を満たす、ベンチマークと動向調査、プログラムの効果と改善の長期的な記録、学生の成績を全国標準と比較、カリキュラムの開発・改善、学生による上級科目の履修資格の判定。
	Major Field Tests (GRE Subject Test)	1990	学生・課程	専門分野の専攻修了時点で習得していることが期待されている概念・原理・知識の習得	学生：総合的科目として、あるいは最終学期に卒業要件としてテストが実施される場合が多い。高等教育機関：ハイスティクス（中～高度）な判断をするために利用する。専門分野における最終学期の達成度を記録して学部カリキュラムの効果を評価する、アクレディテーションおよびアカウンタビリティ要件を満たす、ベンチマークと動向調査、プログラムの効果と改善の長期的な記録、学生の成績を全国標準と比較、カリキュラム開発・改善、学生による上級科目の履修資格の判定。
	CLA (Collegiate Learning Assessment)	2002	課程・機関	実践的作業にもとづく大学レベルの高次思考スキルの評価	高等教育機関：アカウンタビリティ・データを提供する、科目設置に関する全学的な基準を策定する、ベンチマークと動向調査、プログラムの効果と改善の長期的な記録、学生の成績を全国標準と比較、カリキュラムの開発・改善。

<p>米 国 と カ ナ ダ</p> <p>NSSE (National Survey of Student Engagement)</p> <p>米 加 2000 2004</p>	<p>課程・機関</p>	<p>学習の二次的指標： 大学での学習機会への学生参加に関する情報、高等教育機関での経験に関連する学業技能の習得についての自己報告</p>	<p><b>高等教育機関</b>：学生の取り組みを促すためのカリキュラムやサービスの開発・改善、ア krediyteeshon およびアカウンタビリティの要件を満たす（地域ア krediyteeshon における NSSE の活用を促進するために専用キットが用意されている）、ベンチマークと動向調査、プログラムの効果と改善の長期的な記録、学生の成績を全国標準と比較。<b>一般社会</b>：公開集計結果より、学生が高等教育の経験から何を習得しているのかについて情報をえることができる。<b>政府</b>：ア krediyteeshon 指標に採用。データは全国的・機関種別ベンチマークに対応。</p>
<p>GRE (Graduate Record Examination, General Test)</p> <p>1966</p>	<p>学生</p>	<p>累積的学業能力、大学院での研究に必要な一般学業技能の習得</p>	<p><b>高等教育機関</b>：入学希望者の成績を比較して、選抜に関する意思決定の資料とする。</p>

# 高等教育における学習成果アセスメント

- 特筆すべき事例の比較研究 -

OECD 教育関連ワーキングペーパーNo. 15

---

平成 20 (2008) 年 9 月 1 日発行

Deborah Nusche 著

深堀 聰子 訳

〒100-8951 東京都千代田区霞が関 3 丁目 2 番 2 号

国立教育政策研究所

高等教育研究部

TEL 03-6733-6972

E-mail [fukahori@nier.go.jp](mailto:fukahori@nier.go.jp)

URL <http://www.nier.go.jp>

印刷所 (株) 太陽美術

〒135-0024 東京都江東区清澄 2 丁目 7 番 11 号

TEL 03-3642-6045

---