

－ ノート －

規格化ゲインを用いた学習成果の定量化

浅木森 和夫

A Measurement of Learning Outcomes Based on the Normalized Gain

Kazuo ASAKIOMRI

要 旨

文書処理技術を学ぶ「生活情報技術」の授業において、初回の授業で実施した事前テスト（プレテスト）と一定回数授業が終了した後で実施した事後テスト（ポストテスト）の結果を用いて、規格化ゲインによる学習成果の定量化を行った。2013年から2015年の3年間のデータについて、平均規格化ゲインとして 0.49 ± 0.02 という値を得た。

キーワード：Learning Outcomes、Assessment、Normalized Gain
学習成果、評価、規格化ゲイン

1. はじめに

学習者が学習を通して得る成果は、学習者が学習前に持っていた知識や技術を学習を通して更新し、そのことが、学習後の学習者の行動変容につながることによって評価される。

教育は、学習者、教師、教材などいくつかの要素から構成される一つのシステムとしてとらえることができる。学習者が、学習前に持つ知識や技術、態度がシステムへの入力となり、さまざまな教育活動を通して、学習成果として新しい知識や技術、考え方などを得てシステムから出力される。システムからの出力、それが学習成果であるとしてとらえることができる。

そのため、教育システムをよりよく稼働させるためには、教育活動を通して学習成果が一定の基準に達するように、システムの構成要素を調整する必要がある。そこで、インプットとアウトプットの関係を定量的に分析し、システムを改善して行くことが求められる。

システムからのアウトプットとして設定されるものが学習目標である。目標としては、知識の習得、運動能力の向上を通じた技術の習得、学びへの関心や学習への意欲を深める情意に関係するものが設定される。

そして、学習目標が達成できたかを評価するときに用いられるものがさまざまな試験、テストである。テストは、学習の最終段階で実施されることが多い、知識については、学んだ内容の記憶や理解を測定するペーパーテスト、学んだ技術の習得については実技試験が用いられる。

情意分野については、テストを通して学習成果を定量化することはむずかしい。この分野では、質問紙法がよく利用される。この方法では、対象者に自身の属性や心理状態、行動傾向を質問紙によって自己報告式に回答してもらう。測定には、尺度が必要であり尺度の信頼性や妥当性によって得られた数値の確度が決まる。

テスト結果は、数値化することによって情報をうることができる。テスト理論は、数値化に合理性や客観性を与えるものであり、よく知られたものに項目反応理論（Item Response Theory）がある。

Kang^[1] たちは、項目反応理論を用いて、神戸女子短期大学において開講されている「情報リテラシー」で実施しているアンケート調査を用いて受講生の経年的な意識変化を測定した。

その結果、2004年度～2006年度に収集した回答の定量的分析結果に、学生意識に経年的な変化の様相が見られた。この時期は、旧学習指導要領で学んできた学生と新学習指導要領で学習してきた学生とが交錯する時期と一致しており、我が国が進めてきた21世紀の「教える」教育から「学ぶ」教育へのシフトで、「ゆとり」を中心とする意識刺激型の教育が学力低下を招いたという評価・反省との関連において興味ある結果であった。

このように教育におけるアウトプットは、さまざまなテストやアンケートを利用して測定されている。しかし、それらの多くは授業の最後に実施され、点での測定に終わっている。

時間とともに教育は、教えることへの評価から学ぶことへの評価に軸足を移しつつあり、教育の視座が教師から学習者に変化してきている。

学習者の視座では、学習者が持つ学習前能力が、教育活動によってどのように変化しアウトプットされたかが重要であり、アウトプットの評価を、点から線へそして面へと拡大することが求められる。

学習者が、学習前に持っていた先行知識、先行技術そして関心、それと学習目標として設定された目標との定量的な差、言い換えれば、学習者の伸びしろと、学習をとおしてどのように

先行知識や技術、関心が変化したか、学習で得た変化の関係を定量的に測定することが学習成果の情報となる。

すなわち、学習成果であるアウトプットとインプットの差を最適化することによりよりよい教育システムを構築することにつながる。

その一つの試みとして、R.Hake^[2]により物理教育におけるインターアクティブな授業方法の効果を分析するために導入された規格化ゲインを用いて、文書処理技術を学ぶ「生活情報技術」の授業について学習効果の分析を行った。

2. 生活情報技術

「生活情報技術」は、神戸女子短期大学総合生活学科1年生を対象に後期に開講している。この授業は、演習科目として開講されており、文書処理技術を中心に情報社会における仕事へのコンピュータやネットワークの効率的な利用をテーマにしている。

そして、技術目標として、事務仕事で中心となるビジネス文書が作成できることを設定している。具体的な学習成果として、基本的なビジネス文書（案内文）の作成がワードを用いてできることを要求している。

授業は、ビジネス文書の基礎、日本語の基礎、分かりやすい文章作成技法、ワードを用いた文書作成技術、情報社会での仕事の仕方を内容としている。

1回90分の授業では、前半60分を講義にあてている。そして、残りの30分でビジネス文書作成実技試験を実施している。実技試験は、日本商工会議所の日商PC検定文書作成における出題形式を踏襲し、文書作成技巧問題を出題している。^{[3][4]}

また、授業内容の定着を図る目的で、授業開始時に毎回「確認のための小テスト」を実施している。講義トピック終了後にはトピックのまとめを作成、提出させている。

成績評価は、確認のための小テストの総合成績、授業トピックのまとめ、文書作成実技試験、そして、最後の授業で実施する総まとめの知識・技術試験の結果を総合して行っている。

学生の中には1年生前期に開講されているこの授業の内容と関係する科目を履修しているものもいる。そこで、最初の授業で前期この授業の内容と関係する科目を受講したか、また、この授業を受講する目的などについてアンケートを行っている。そして、アンケート結果は授業進行に反映させている。

なお、今回の分析にはアンケート回答結果を使用していない。

3. 規格化ゲイン

近年さまざまところで、教師が一方的に話をする伝統的な授業と学習者が中心となって学習に取り組むインターアクティブな授業について、その効果についての比較研究が行われている。

R. Hake^[2] は、「入門物理学」授業において伝統的な授業方法とインターアクティブな授業方法について授業効果を規格化ゲインを用いて分析した。

規格化ゲイン g は

$$g=(S_f-S_i)/(100-S_i)$$

で定義される。

ここで、 S_i と S_f は教育効果を測定したいコースにおいて授業期間の最初と最後に実施するプレテスト、ポストテストのクラス平均正答率である。なお、測定にはプレテスト、ポストテストとして同じ問題が用いられる。

もし、ポストテストの平均正答率が $S_f=100$ の場合には規格化ゲインの値は1となる。また、ポストテストの平均正答率が、プレテストの平均正答率と変わらない場合は分子が0となり、ゲインも0となる。

このように、規格化ゲインは0～1の間をとり、ゲインが1に近いほど学習成果があったものと考えられる。

Hake は、約6,000人のさまざまな高校生、大学生について伝統的な授業を受けた学習者とインターアクティブな授業を受けた学習者のデータをもちいて規格化ゲインを計算し、伝統的な授業を受けた学習者について 0.23 ± 0.04 、インターアクティブな授業を受けた学習者については 0.48 ± 0.14 という結果をえた。

この結果は、インターアクティブな授業が教育効果をうるには有効であることを示した。

石本たち^[3]は、基礎力学の授業を通して、学習者が物理事象を物理学的にどの程度とらえられたか規格化ゲインを用いて分析した。その結果、日本においても修得度を知る上で規格化ゲインが有効であることを示した。

4. 規格化ゲインを用いた授業成果の分析

「生活情報技術」を受講する学生が、前期、「情報リテラシー」を履修したり、ビジネス関連科目を履修しておれば、ワードの利用技術やビジネス文書の基本的な構造、書き方について先行知識を有しているものと予想される。

そこで、規格化ゲインを用いる授業成果の調査として、技術目標としているビジネス文書の基本となる案内文の作成を用いることとした。

プレテストとして、最初の授業において、案内文作成の実技試験を実施した。試験では、問題文に条件・背景を示し、ワードを用いて示された条件・背景にあう適切な社外ビジネス文書を作成させた。ビジネス文書のレイアウトや文書表現は知識であり、ワードを用いた文書入力には技術と考えられる。

制限時間は約30分、解答は授業使用しているラーニングマネジメントシステム（Moodle）のファイルアップロード機能を用いて提出させた。

なお、プレテストでは、テストの目的は先行知識や技術を把握することであり、結果を成績評価には含めないことを説明している。

プレテストの評価は、ビジネス文書の主要項目が適切に表現、配置されているか、文章の表現が適切か、について行った。これらは、今後この授業で詳しく学ぶ事柄である。

ポストテストは、授業の約3分の2が終了した10回ないし11回目に内容がプレテストと同じであることを説明せず実施した。問題、採点についてはプレテストと同じものを用いている。

なお、このテストの成績が通常の文書作成実技試験と同様に成績評価に含まれることを説明した。

4.1 規格化ゲイン（全体）

2013年から2015年まで、3年間の結果を表に示す。表には、分析に使ったデータの数、100点を満点としたプレテストの平均、ポストテストの平均、そして、それを用いて計算した規格化ゲインである。なお、誤差は統計的誤差を示している。

表1 規格化ゲイン

年度	人数	プレテスト平均	ポストテスト平均	規格化ゲイン
2013	18	23.91±2.49	60.58±2.14	0.48±0.02
2014	9	43.14±6.04	78.43±3.40	0.62±0.04
2015	14	45.00±2.67	65.00±3.70	0.36±0.04

年度ごとのデータ数、プレテスト平均、ポストテスト平均、規格化ゲイン

年度によって、規格化ゲインに統計的誤差を超えて変動がみられる。その理由として、年度ごとの学習集団の特性が影響しているものと考えられる。3年間の規格化ゲインの最確値を求めると0.49±0.02であった。

4.2 受講者のそれぞれの規格化ゲイン

年による規格化ゲインのふらつきは、統計的ふらつきで説明することはむずかしい。この原因は、年度ごとの履修者の特性によるものと考えられる。

そこで、学習者個々について考えてみる。学習者個々の規格化ゲインを次のようにして求める。

学習者のテスト結果を、100点満点として換算する。その成績を基に「100－プレテスト成績」を求める。この値は、学習者の伸びしろを表している。次に、ポストテスト終了後、成績を100点満点に換算して「ポストテスト成績－プレテスト成績」を求める。この値は、学習の成

果を表している。そして、この値を伸びしろで割って学習者の規格化ゲインとする。

例えば、ある学習者についてプレテストの成績が60点、ポストテストの成績が80点の場合の場合、伸びしろは、 $100 - 60 = 40$ 点、学習成果は $80 - 60 = 20$ 点となる。したがって、この学習者のゲインは、 $20 / 40 = 0.5$ となる。

図1に3年間のデータについて規格化ゲインの分布を示した。規格化ゲインの平均値は0.47、標準偏差は0.15であった。

分布をみると規格化ゲインの値が0.3を下回るもの、0.7を上回るものが少なからず存在する。

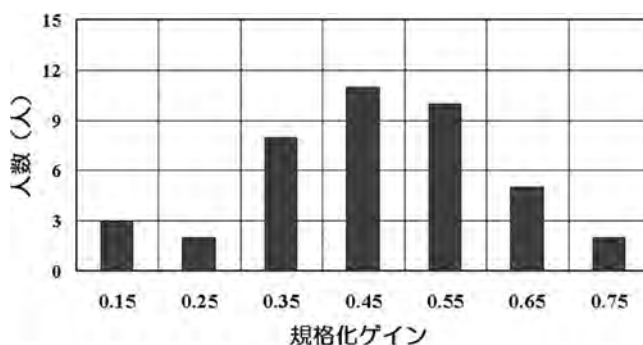


図1 規格化ゲインの分布

規格化ゲインは、学習者の授業前に持つ伸びしろと学習によって得られる学習成果に関係するため、個々の学習者の伸びしろと規格化ゲインの関係について分析することによって、履修者の特性を知ることができる。

図2には、プレテストとポストテストの相関を示した。図では、年度ごとにマークを変えている。点線は、規格化ゲインが0と0.5のラインを示している。

プレテストの成績が30~40付近に規格化ゲインの低いデータがみられる。

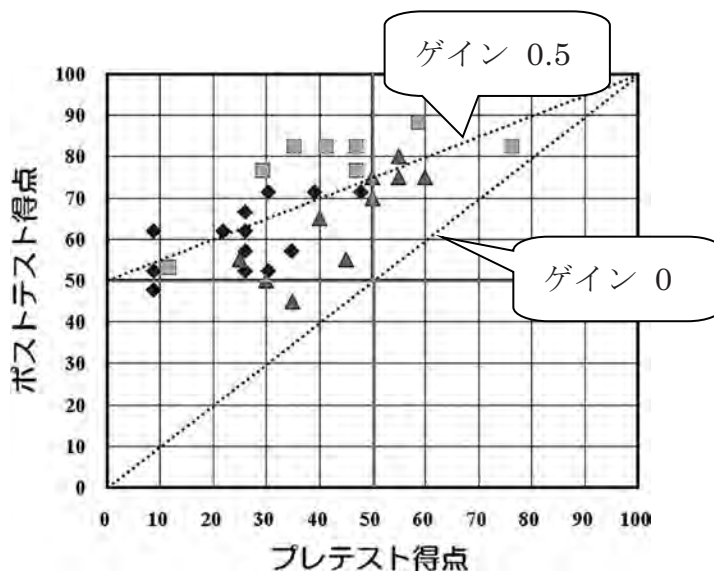


図2 プレテスト成績とポストテスト成績の関係
 ◆：2013年 ■：2014年 ▲：2015年 点線はゲイン0と0.5のラインを示している

図3には、伸びしろとゲインの関係を示した。授業開始時にある程度の先行知識や技術を有しているときには伸びしろは小さくなる。そのため、伸びしろの少ない学習者が、規格化ゲインを上げるためには、少なからず学習努力が要求される。しかし、学習者の伸びしろが小さい、プレテストの成績がいいことは、学習者が先行知識として学習目標に近い能力を有していることを示している。そのため、このような学習者の規格化ゲインが小さな値であっても、学習目標という意味では影響は小さい。

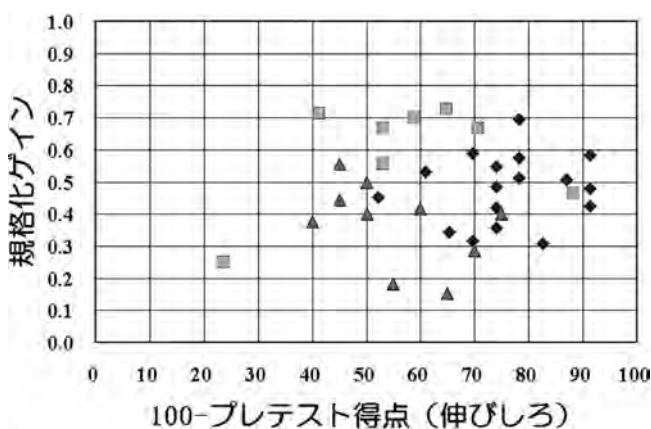


図3 伸びしろと規格化ゲインの関係
 ◆：2013年 ■：2014年 ▲：2015年

しかし、伸びしろの大きな学習者が小さな規格化ゲインを持つときには問題となる。すなわ

ち、このような学習者は、伸び幅が大きいのに学習を通して力を伸ばすことができていないことを示している。

その原因として、

- ・ 先行知識や技術が極端に乏しく、最初から授業についていくには難があった
- ・ 授業に関心が乏しく、積極的に学習に係わることができなかった

などが考えられる。

5. まとめ

授業成果を定量化するための一つの試みとして「生活情報技術」の授業において、規格化ゲインを導入して学習成果について評価を行った。

授業目標である「ビジネス文書作成実技試験」を用いた学習成果の検証において、3年間のデータから得られた規格化ゲインは 0.49 ± 0.02 であった。また、データには年度ごとに統計誤差を超える変動がみられた。年度ごとの変動の原因としては、受講生の特性が年度で変化していることが考えられた。

規格化ゲインは、先行知識と学習目標の間に大きな差、学習者が持つ学習の伸びしろと学習によって得られた学習成果が関係する。大きな伸びしろを持つ学習者は、積極的な学習を通して新しい知識や技術の習得が可能である。そのため、そのような学習者にとって規格化ゲインは、学習成果のいい指標になっており、学習の程度を定量的に相対化することができた。

また、伸びしろをパラメータに規格化ゲインを調べることにより、学習者が学習前の先行の能力として授業を受講するにふさわしい能力を有していたか、学習者が積極的（アクティブ）であったか、非積極的（インアクティブ）であったか等、学習者の状況について情報を得ることができる可能性が見られた。

精度よい規格化ゲインが得られるためには、基準テストの開発を進めることが中心的な課題である。

また、学習者個々の規格化ゲインを指標として、授業で課する課題やテストの成績、アンケー

ト回答結果などとの相関を分析することにより学習者の多面的な学習成果を客観的に知るための仕組みを確立することができると予想される。

参考文献

- [1] Kanetsugu Sakata, Min Kang, Kazuo Asakimori and Makoto Kaburagi, “An Analysis for Student Opinion Survey with Item Response Theory” Proceedings 8th International Conference on Information technology Based Higher Education and Training, (2007) 403
- [2] Richard R. Hake “Interactiv-engagement versus traditional method: A six-thousand-student survey of mechanics test data for introductory physics courses”, Am. J. Phys. 66(1998)64.
- [3] 石本美智、木村正廣 “物理授業改善のための学習効果測定試験” 高知工科大学紀要 (2005) 218.
- [4] 日本商工会議所編 “ネット社会のデジタル仕事術” 廣濟堂出版 (2006)
- [5] 日本商工会議所編 “日商 PC 検定試験 3 級公式テキスト文書作成” FOM 出版 (2006)