

# 「GPAS法」によるオンライン理解度テストの有効性評価 - その授業改善への効果 -

## Assessment of Online Comprehension Test by “GPAS Method” - Its Effective Feedback to the Teaching Improvement -

小林 浩 佐野 香 中村尚五  
東京電機大学情報環境学部

Abstract: Methods to evaluate effectiveness of ICT-applied teaching have not established yet. The GPAS method, previously proposed by the authors, measures and utilizes the differences in the degree of ICT-utilization among individual students, and can be called “individual difference method”. The idea of this method is to classify the students into several groups according to their degree of perseverance, while influential factors related to the potentialities of students, and to peculiar to the course were eliminated by using GPA. The average values of the degrees of ICT-utilization and perseverance were statistically tested for significant difference between the two end-groups. The feedback on the evaluation results of the online comprehension test was given to the author’s class, which resulted in about remarkable improvements in the class.

Keywords: ICT-applied class, Educational Technology, GPA, effectiveness evaluation, GPAS method

### 1. はじめに

近年のパーソナルコンピュータやインターネットの急速な普及とともに、学校教育の場においてもe-Learningや講義ビデオなど様々なICT (Information and Communication Technology, 以下の本文中では単に“IT”と略す)の活用が図られるようになってきた。しかしながら、その有効性を迅速かつ適切に把握し、授業改善にフィードバックする方法は必ずしも確立している状況にはない。

すなわち、一般的に新しい手法や新薬などの有効性評価は、新手法などの適用前の定常状態における成績と適用後のそれとを比較する方法と、対象を適用組と非適用組とに分けて比較する方法とに大別される。ここでは、前者を時間差法、後者を差別化法と呼ぶこととする。前者の代表的な方法として、プリテ

ストとポストテストの得点の相関性を利用して学習効果を回帰成就値で表す方法<sup>[1]</sup>があるが、高校の英数国のような教科には有効なもの、学生たちが白紙の状態で臨む専門科目には適用し難く、またフィードバックのために評価結果を学生たちに直感的に理解させることも難しい。一方、差別化法は同じ内容の授業が複数の異なる学校で開講されているケースには適用しやすいが、年度単位でしかも1授業しか開講されない大学などの専門科目には公平性の点から適用し難い。

こうした中、時間差法と差別化法とを組み合わせ合わせた適用事例として、全国10箇所に散在する専門学校で10科目を対象に衛星通信とCAIとを併用した遠隔授業と、各校独自の座学授業間で成績を比較評価した報告が中山らによってなされている<sup>[2]</sup>。同報告によれば、全体的に遠隔授業のほうが成績は高めであったが、これは衛星を介して授業が放映されるため、講師は通常の授業より数倍の時間をか

Hiroshi Kobayashi\*, Kaoru Sano and Syogo Nakamura  
Tokyo Denki University  
\*E-mail: hiroshi\_kobayashi@sie.dendai.ac.jp

けて準備したことなどが寄与したようである。ただし、CAIについては講師へのインタビューなどをもとに、効果は大きくないとの言及に止まっている。また、牟田らによる計5校の専門学校におけるCAIの有効性評価報告でも、顕著な効果は見出せなかったとしている<sup>[3]</sup>。これは、測定値（成績）をクラス全体で平均化、すなわちCAIを積極的に活用した学生と積極的にでなかった学生とを混在したまま評価したため、その有効性が全体の中で薄められたことが要因の一つとして考えられる。

GPAS (Grouping in the Plane of Accumulated Grade Point Average and normalized Score) 法は、学生ごとのIT活用度合の違いを利用してその有効性を評価しようとする個体差法と呼べるものである。そのポイントは、総合成績評価指標として国際規格のGPAを用いて学生の総合能力などの影響要因を隠蔽してから、学生たちの頑張り度合によってグルーピングし、グループ間でIT活用度合と頑張り度合の相関性を調べることにある<sup>[4]</sup>。

筆者らの担当科目「インターネット総論」では、IT活用の一つとして教科書の章単位でオンライン理解度テストを実施しているが、中間試験直後、GPAS法を用いて同テストの有用性を学生たちに訴えたところ、期末試験後の調査で、理解度テストに積極的に取り組むようになった学生が大幅に増え、また得点（素点）ベースの平均点も大きく向上するなど、授業改善に顕著な効果が認められた。

以下、GPAS法の概要と、これを用いたオンライン理解度テストの有効性評価、さらに同評価結果の授業へのフィードバックによる改善効果について報告する。

## 2. GPAS評価法

GPAS法は科目に固有な特性（試験成績点の平均や偏りなど）や個人レベルの影響要因の隠蔽、授業フィードバックにおける学生たちへの説明の明快性・理解性などに留意して開発されたもので、その骨子は次の通りである。

学生  $k$  の中間試験などの得点を式(1)により規格化評点  $NS_k$  に変換し、点  $(NS_k, GPA_k)$  を  $NS$ - $GPA$  平面上にプロットする(図1)。

表1の各評価ポイント  $GP$  と、これに対応する評点の下限値を結んだ  $NS$ - $GPA$  直線を上記平面上に描き、同直線から点  $(NS_k, GPA_k)$  までの水平方向距離を学生  $k$  の頑張り度合と定義する(図1)。

頑張り度合によって学生たちをグループに分け、各グループの平均値を算出する。

アンケート結果やログデータから、学生ごとのIT活用度合を求め、上記各グループの平均値を算出する。

両端2グループ間の平均値の差の検定により有意性が認められ、かつIT活用度合の増加が頑張り度合を増す方向にあれば、当該IT活用は有効であったと判断する。

表1 GPAの定義

評点	評価	評価ポイント(GP)
90以上	S	4
80以上90未満	A	4
70以上80未満	B	3
60以上70未満	C	2
40以上60未満	D	0
40未満	E	0

$$GPA(\text{通算}) = \frac{(\text{各科目の単位数} \times \text{得た評価ポイント}) \text{の合計}}{\text{履修登録した全科目の総単位数}}$$

$$\begin{aligned} & \text{if } S_k > S_c \\ & \text{then } NS_k = (S_k - S_c) \times \frac{NS_{max} - NS_c}{S_{max} - S_c} + NS_c \\ & \text{else } NS_k = (S_k - S_c) \times \frac{NS_{min} - NS_c}{S_{min} - S_c} + NS_c \end{aligned} \quad \text{式(1)}$$

ただし、 $S_k$ :k君の試験得点、 $NS_k$ :同規格化評点(NS)  
 $S_{max}$ :全学生の得点最大値、 $S_c$ :同平均値、 $S_{min}$ :同最小値  
 $NS_c$ :履修登録した全学生のGPA平均値から逆算されるNS  
 $NS_{max}$ :NS最大値(=100:所与)、 $NS_{min}$ :NS最小値(=0:所与)

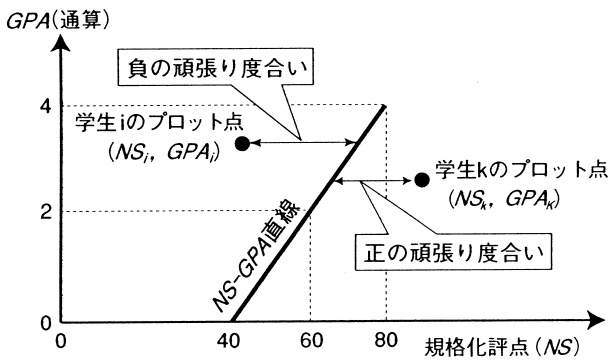


図1 頑張り度合の定義

①の規格化評点操作は、履修登録した学生全員の前 Semester までのGPAの平均値に対応する評点平均値 $NS_c$ に、試験得点の平均値 $S_c$ を移動させ、かつ0~100点まで広く分布させることによって、科目固有の特性を隠蔽しようとするものである。

ところで、新たに履修する科目の成績は英語や数学などのいわゆる5教科の総合成績との関連性をもつという前述の中山らの報告<sup>12)</sup>や、大学専門科目を対象とした岡田らの報告<sup>15)</sup>は、試験の合成効果<sup>14)</sup>を反映したものと考えられる。

②は、この効果を利用して個人レベルの影響要因(総合能力)を隠蔽しようとするもので、図1に示すNS-GPA直線より右側に位置する学生kは $GPA_k$ から期待される成績より良かった(頑張った)、逆に左側の学生iは怠慢だったと解釈できる。

③から⑤は、例えば各グループ内の学生数が同じになるように3組に分け、両端2グル

ープ間の統計量(平均値)に相関性が認められるか否かを判定しようとするものである。

### 3. オンライン理解度テストの有効性評価

「インターネット総論」では、計1,200項目の用語の理解と獲得を求めているが、学生たちが効率よく学習できるよう教材の配布や講義ビデオ配信など、様々なITの活用が行われている。これらの中で、最も学生たちの活用度合が高く、またその効果が大きいと期待できるのがオンライン理解度テストである。これは、学生たちに復習を定着させることを狙ったもので、章単位で空所5~10の穴埋め問題を25問程度準備しておき、その中からランダムに5問を出題している。また、用語やプログラムの一部を記述させたり、ランダムに配列された選択肢から選ばせたりするなど、図表を交えた多彩な出題が可能である。

オンライン理解度テストの得点は、履修科目の評点にいくばくか上乘せすることにしてあるので、学生たちは満点を取るまで何回もトライしているが、単に(答えを書き取り)点数をとるためだけにアクセスしている学生や、十分理解し納得するまで時間をかけて解答に取り組んでいる学生など、その活用度合は大きく異なっている。

図2は、ほぼ同じ試験構造の中間試験と期末試験(空所計60の穴埋め問題6問と、論述問題4問から構成)について、上述の手順に沿って学生たちの試験得点を規格化し、NS-GPA平面上へプロットしたものである。各プロット点の記号は理解度テストに関する取り組み姿勢を示すもので、各試験終了直後に実施したオンラインアンケート(三者択一)で学生たちが回答した主観データを用いている。これは、オンライン理解度テスト用に市

販のe-Learningシステムを用いたため、所望のデータを取得できなかったことによるもので、システムのログデータ（受験回数や受験時間など）などに基づく客観データの取得により、客観性を高める余地は残っている。なお、学生たちの Semester 内での履修離脱による統計量のゆがみを避けるため、中間および期末試験両方を受験し、かつアンケートに回答した計111名のみを統計対象とした。

また、NS-GPA直線を平行移動して各グループの標本数が同数となるよう境界線を設定した。最も頑張った学生が多いグループを頑張り組、以下順に実力相応組、怠慢組と命名している。同図から、中間試験ではプロット点が広く分布し、かつアンケート結果も多彩であることが、逆に期末試験ではプロット点が境界線付近に集まり、納得するまで何度も取り組んだ学生が増え、答え合わせに終始した学生が激減していることが分かる。

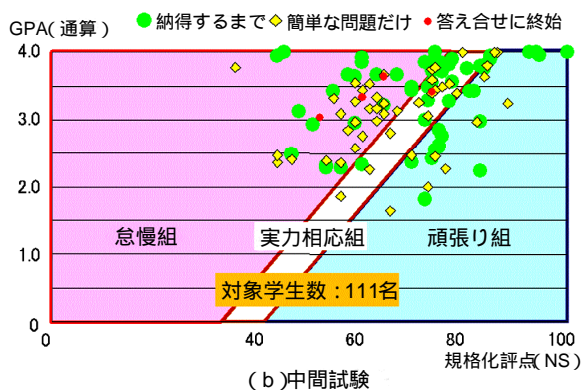
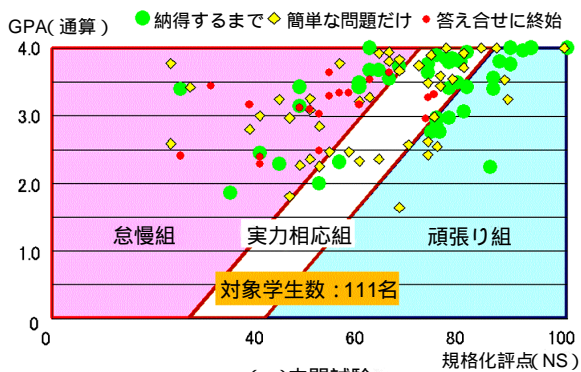


図2 NS-GPA平面へのプロットとグルーピング

また、表2は頑張り組と怠慢組について、IT活用度合と頑張り度合の平均値と、その差の検定における有意水準を示している。ただし、IT活用度合は前述の理解度テストに関する取り組み姿勢を順位尺度（納得するまで取り組んだ：1，簡単な問題だけ解いた：0，答え合わせに終始した：-1）として表し、さらに隣接する値に等間隔を想定して間隔尺度としてt検定に用いた。これは、心理学などのように定量的なデータを取得し難い分野で便宜的に用いられている方法である。

中間および期末試験とも、二つの度合が高い有意差を示し、かつIT活用の増加が頑張り度合を増す方向であったことから、この科目では Semester を通してIT活用が有効であったと評価される。特に、中間試験では両方の度合とも有意水準は1%以下であったことから、IT活用効果は極めて高かったと判断される。なお、期末試験では有意水準が3%以下となったが、これはアンケート順位尺度-1を回答した学生が激減し、2グループ間の平均値の差が縮まったためである。

表2 両端2グループ間の平均値の差の検定

(a) 中間試験(自由度: 36)

グループ	IT活用度合の平均値	頑張り度合の平均値
頑張り組	0.51	8.0
怠慢組	-0.03	-25.0
有意水準	1%以下	1%以下

(b) 期末試験(自由度: 36)

グループ	IT活用度合の平均値	頑張り度合の平均値
頑張り組	0.65	9.0
怠慢組	0.35	-16.6
有意水準	3%以下	1%以下

#### 4. フィードバックによる授業改善効果

上述したように中間試験を挟んでIT活用度合などに大きな変化が観測された。これは、

学生たちが中間試験の成績の悪さにショックを受けたことと相俟って、中間試験直後の授業で図2(a)と表2(a)などを提示しながらオンライン理解度テストの有用性を訴え、奮起を促したことが大きな契機となったためと思われる。

すなわち表3に集約されるように、GPAS法による評価結果を学生にフィードバックしたことによって、中間試験後、理解度テストに真面目に取り組むようになった学生の大幅な増加とクラス全体の学習意欲の向上（頑張り度合分布の縮小）をまねき、その結果が成績の向上と成績格差の是正に反映されたものと判断される。

なお、GPAと規格化評点間の相関係数は期末試験のほうが小さく、また同試験ではGPAと頑張り度合とが負の相関を示していることから、成績下位（GPAが低かった）学生の改善により効果的であったと解釈できる。詳細は省略するが、これらは前述の回帰成就値による評価と意は同じである。

表3 中間/期末試験結果に見られる改善効果

	中間試験	期末試験	改善率	
納得するまで取り組んだ学生の構成比率	39.6%	52.3%	+32%	
頑張り度合の標準偏差	15.3	11.7	+24%	
GPAと規格化評点との相関係数	0.45	0.35	-	
GPAと頑張り度合との相関係数	0.12	-0.12	-	
(参考)	試験得点の平均値	37.6	52.8	+41%
	試験得点の標準偏差	15.0	12.7	+15%

## 5. おわりに

学生ごとのIT活用度合の違いを利用したGPAS法の紹介と、これを用いたオンライン理解度テストの有効性評価、またこの評価結

果の学生たちへのフィードバックが Semester 内での授業改善の契機として効果的だったことを報告した。

GPAS法は授業科目や対象とするITに依存することなく広く適用できる。また、これまで試験成績という客観的測定に止まりがちであった学校教育において、学生たちに頑張り度合を明示して公平に評価する、すなわちステップアップの動機付けを行っていく手段として活用していくことも期待されよう。

今回はオンライン理解度テストのみに着目したが、複数のITを活用する授業に対応できるようにするには、多変量解析などとの併用による拡張が必要である。加えて、長期的な課題として、本手法や授業コンテンツなどとの連携により、学生たちの学習意欲と効率を高めるような新しい概念の学習進度評価システムの研究に発展させていく所存である。

## 参考文献

- [1] 池田央: テストと測定. 教育学大全集, 第一法規出版, 1982.
- [2] 中山実, 清水康敬: 通信衛星による講義とCAIを併用する遠隔教育システム(PINE-NET)の学習成績による評価. 日本教育工学雑誌, Vol.17, No.2, pp.85-92, 1993.
- [3] 牟田博光, 星野敦子, 坂元昂, 奥山明: CAROLシステムの利用による教育効果の評価研究. CAI学会誌, Vol.8, No.2, pp.59-69, 1991.
- [4] 佐野香, 小林浩, 中村尚五: 通算GPA-規格化評点平面上でのグルーピングによるIT活用授業の有効性評価. 信学技報, ET2003-67, pp. 41-46, 2003.
- [5] 岡田幸子, 菅野文友: 電子計算機演習の成績に影響を与える学生の諸因子について(教育データによる解析事例). 情処研報, コンピュータと教育8-3, pp.1-8, 1989.