

【原著】

新潟大学の個別学力検査における 作題技術の向上を目的とした 作題者へのフィードバック内容の整理

佐藤喜一，並川 努，濱口 哲（新潟大学）

本稿では、個別学力検査における作題の PDCA サイクル（plan-do-check-act cycle）構築に向け、入学センターで検討した個別学力検査の作題者へのフィードバック内容について分析例を提示した。今回の試みでは、大問得点のみを活用し、試験終了後に作題者へフィードバックする内容として、試験の構成、受験者、得点の分布、学部別の得点分布、信頼性と項目分析、相関係数と共分散比をリストアップした。

1 背景と目的

新潟大学（2010）では、第 2 期中期目標・中期計画（平成 22～27 年度）の一つとして、「アドミッションポリシーを明確にして、それに即した入学者選抜方法となるよう、入学者選抜制度の充実を図る」という目標を掲げている。その目標の達成に向けた具体的方策の一つとして、「入学試験問題の改善を図るために、入学センターで、試験問題の妥当性、信頼性等の検証を行う」ことが計画されている。

この計画について、著者らは、入学センターにおいて個別学力検査の試験結果を検証するととどまらず、作題者にも試験結果の統計情報をフィードバックして試験問題作成の PDCA サイクル（plan-do-check-act cycle）を構築したいと考えている。入学試験問題の改善のためには、大学側の作題技術の向上が一つの鍵であり、作題の PDCA サイクルを構築することによって継続的な作題技術の向上が期待できる。

本稿では、個別学力検査における作題の PDCA サイクル構築に向け、入学センターで検討した個別学力検査の作題者へのフィードバック内容について分析例を提示する。

2 本学の個別学力検査

本学では、国語、地理歴史、公民、数学、理科、外国語、総合問題、実技、小論文、面接・口頭試問を組み合わせた個別学力検査が募集単位ごとに実施されている。このうち、分析対象となる筆記試験のほとんどは 3, 4 問程度の大問から構成されており、さらに大問は関連する数個の小問から構成されている。成績データとしては大問得点と小問得点が得られるものの、今回の試みでは大問得点のみを活用して作題技術の向上に役立つような統計情報を作題者へフィードバックすることを想定した。

本学は九つの学部をもつ総合大学であるため、個別学力検査の利用のされ方がかなり複雑な教科も存在する。表 1 の「物理」は複雑な場合の例であり、学部（学科）によって受験を指定される問題番号の組み合わせが多岐にわたる。この場合、すべての受験パターンの分析結果を作題者にフィードバックしても、テストの専門家ではない作題者にとって、その内容を理解した上で次回の作題に活かすことは容易なことではない。それゆえ、入学センターでは、たとえすべての大問を分析できなくても、もっとも代表的な受験パターンに

ついでのみ分析結果をフィードバックすることにした。

次節では、本学の個別学力検査における試験結果の一部を利用し、入学センターで検討した個別学力検査の作題者へのフィードバック内容について分析例を提示する。その際、肥田野（1972）、池田（1973）、林（2001）、適性試験管理委員会（2011）、適性試験委員会（2011）を参考資料とした。

表1 平成23年度「物理」の試験構成

志望学部(学科)	問題番号			
	①	②	③	④
教育学部	○		○	○
理学部(物理学科)	○	○	○	○
理学部(数学科・生物学科・地質科学科・自然環境科学科)	○	○		○
医学部	○	○	○	
歯学部	○	○	○	
工学部	○	○		○
農学部	○	○		○

3 フィードバック内容

3.1 試験の構成

前述の通り、作題者の理解しやすさへの配慮から、もっとも代表的な受験パターンについてのみ分析結果をフィードバックする。作題者には、最初にどの受験パターンを分析したのかをフィードバックする。

本稿の分析例として、単純な試験構成の例を表2に示す。この場合、もっとも典型的な受験パターンは、受験者数が突出した大問1, 2, 3の組み合わせである。黒枠は、A, B, C, D学部の試験結果を利用し、大問1, 2, 3の受験パターンについて分析したことを示している。

3.2 受験者

作題する上で、どんな受験者が何名くらい受験するかを把握しておくことは大切なことである。それらは、出題レベルや試験時間の

設定、採点の労力を考慮した問題形式の選択などに影響を与えるからである。作題者には、受験者数と受験者の属性についてフィードバックする。

具体的には、分析した受験パターンにおける学部別の受験者数（表3）と受験者の学部別比率（図1）をフィードバックする。前節の試験構成と合わせ、作題者は担当教科について典型的な受験パターンと受験状況の実際を把握することができる。

表2 分析された試験構成

	①	②	③	④	受験者数
A学部	○	○	○		250
B学部	○	○	○		250
C学部	○	○	○		250
D学部	○	○	○		250
E学部	○	○	○	○	25

表3 学部別の受験者数

	受験者数
A学部	250
B学部	250
C学部	250
D学部	250
合計	1,000

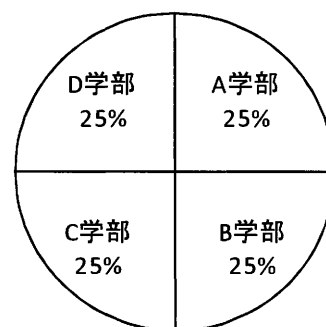


図1 受験者の学部別比率

3.3 得点の分布

テストの得点は、受験者の学力を表現する尺度値であり、受験者や作題者がもっとも関心を寄せる情報の一つである。ここでは、大問得点および合計得点についての基礎統計量（表 4）と得点分布のヒストグラム（図 2, 3）をフィードバックする。

表 4 の基礎統計量から、作題者が想定した出題レベルと実際の出題レベルとの一致の程度を確認できる。作題者の多くは、平均点を 60 点にしようなどと、出題レベルを想定しながら作問するであろう。しかし、作題者が想定した平均点と実際の平均点を一致させることは容易なことではない。表 4 は、問題の難易度についての作題者の認識のずれを修正するのに役立つ。

図 2 と図 3 のヒストグラムをみると、実際の得点分布を視覚的に捉えられる。たとえば、図 2 からは、0-5 点が続出するほど大問 2 が受験者にとって難しい問題であったことや、多くの受験者が 50%未満の低い得点率であったことがわかる。受験者にとって難しい問題は、学力レベルの低い受験者を識別するには向かないものの、医歯学系の受験者のような学力レベルの高い受験者を識別するには有効な問題である。表 4 のみでは大問 2 は相対的に難しかったということ以外は明確にわからず、基礎統計量とともにヒストグラムを確認することは情報の補完に役立つ。

表 4 基礎統計量

	大問 1	大問 2	大問 3	合計
配点	30	30	40	100
平均	16.2	9.7	19.3	45.2
標準偏差	5.0	6.7	6.2	14.5
最小	0	0	0	0
最大	28	27	34	84

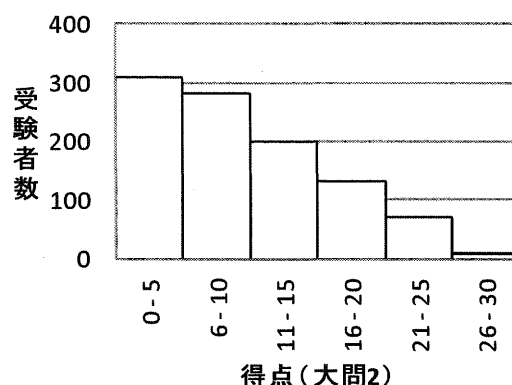


図 2 得点分布 (大問 2)

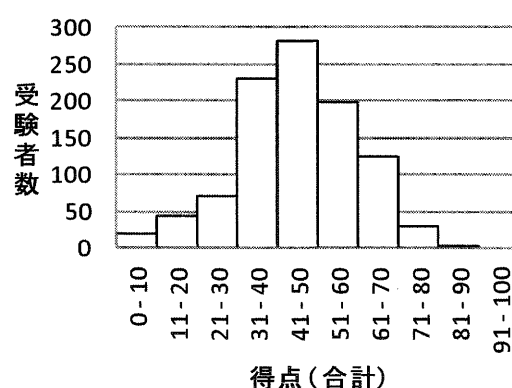


図 3 得点分布 (合計)

3.4 学部別の得点分布

同一構成のテストを複数の募集単位で利用する場合、各募集単位における受験者の学力レベルに注意する必要がある。受験者の学力レベルにあった難易度の問題が出題されないと、受験者の学力を識別しにくくなり、テスト全体の測定精度が下がってしまう。たとえば、学力レベルの高い集団に易しい問題ばかり出題しても、ほとんどの受験者が満点に近い得点をとるため、受験者の学力を精度よく測定できない。とくに、受験者集団に医歯学系の学部が含まれる場合、他学部の受験者に比べて学力レベルが高い傾向があるので注意が必要である。

学部別の学力レベルを作題者が把握できるように、学部別の基礎統計量（表 5）と学部別の得点分布（図 4, 5）をフィードバックす

る。表 5 の合計得点をみると、この場合は C 学部の学力レベルが相対的に高いことがわかる。また、図 4 のような各大問の箱ひげ図を比較すれば、とくに学部間で得点差が生じた大問番号を確認することもできる。各学部の受験者の学力を識別可能なテストになっていたかどうかは、表 5 の基礎統計量と次節の大問得点率分析図を組み合わせることによって確認できる。

表 5 学部別の基礎統計量

	大問 2		合計	
	平均	標準偏差	平均	標準偏差
A 学部	9.8	5.4	46.1	12.1
B 学部	8.3	5.4	39.6	12.7
C 学部	15.4	6.8	55.7	14.4
D 学部	5.4	4.4	39.2	12.4

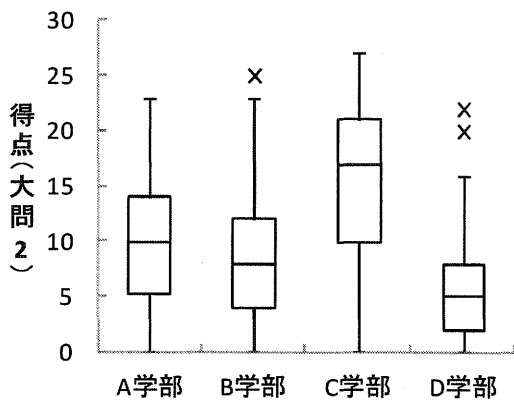


図 4 学部別の得点分布 (大問 2)

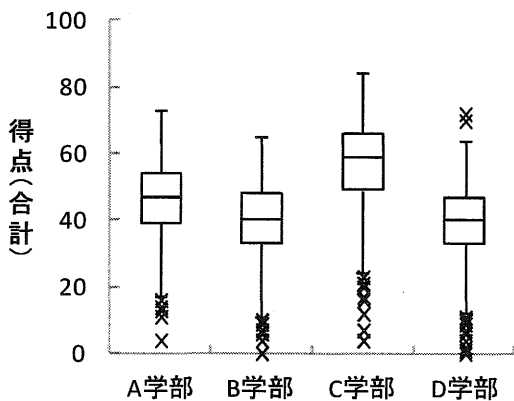


図 5 学部別の得点分布 (合計)

3.5 信頼性と項目分析

試験問題が入学試験としてふさわしい性能をもっていたかどうかを確認することは、試験結果の検証というだけでなく、作題技術の向上にも役立つことである。ここでは、古典的テスト理論に基づき、テスト全体の信頼性を検討するとともに、統計学的にみた個々の項目(大問)の良否についても分析する。作題者には、テストの信頼性(表 6)、大問の特性(表 7)、大問の得点率についての分析結果(図 6、表 8)をフィードバックする。

表 6 に示すように、テストの信頼性は Cronbach の α 係数を用いて記述することにした。Cronbach の α 係数は、テストの内部一貫性の指標であり、テストに含まれる項目群が一貫した特性を測定しているかどうかを表す。その値は、0~1 の範囲を取り、1 に近いほど内部一貫性が高いテストといえる。たとえば、測定したい学力とやや異なる学力を測定する問題がテストに含まれる場合、そのテストの内部一貫性は低くなり、Cronbach の α 係数はその分だけ小さくなる。入学試験は、受験者の進路に大きな影響を与える試験であり、それ相応の高い信頼性が要求される試験である。本学の個別学力検査の場合、3、4 問程度の大問得点のみを計算に利用していることを考慮に入れると、Cronbach の α 係数は 0.6 程度以上の値が望ましいと考えられる。

表 7 には、各大問の特性を表す指標として(平均)得点率、識別力、削除時 α がまとめられている。テストの難易度に関わる側面として得点率がある。得点率(0~100%)は、配点に占める得点の割合であり、その値が高いほど受験者にとって易しい問題であることを示す。得点率の目安としては、難しすぎず易しすぎない 20~80%程度であることが望ましいとされる。また、テストの測定精度に関わる側面として識別力がある。ここでの識別力(通常は 0~1)は、受験者の大問得点と

その大問を除いた合計得点との相関係数であり、その値が大きいほど受験者の学力の識別に資する問題であることを示す。この場合の識別力の望ましい値については、著者らの知るかぎり明確な基準は存在せず、分析結果の蓄積から経験則を導くしかない。現状では、他の大問と比較して極端に識別力が小さい大問は要検討と考えるのが現実的である。さらに、表 7 の削除時 α は、テストからその項目を削除したときの Cronbach の α 係数を表している。削除時 α の値が表 6 の値より大きくなる場合、その大問はテスト全体とはやや異なる学力を測定した可能性がある。いずれの指標においても、望ましい基準から外れた大問については、出題内容も含めた検討が必要である。

各大問の特性をさらに詳しく把握するには、大問得点率分析図（図 6）が便利である。図 6 は、各群の受験者数ができるべく等しくなるように受験者を合計得点の昇順に 5 群に区分し（表 8）、その群ごとに各大問の得点率をプロットしたものである。ある問題が良問であるためには、適度の得点率と一定程度以上の識別力をもつことが必要である。大問得点率分析図を用いる場合、最高でも各群の得点率が 60%程度に達しない問題は難問に分類される。また、学力の高い群ほど得点率は高くなるのが望ましく、その場合にグラフは右上がりの曲線になる。その際、配点の影響はあるものの、一般には識別力が高い項目ほど右上がりの傾きが急になる傾向がある。それゆえ、得点率の変化から必要以上に難問と考えられる場合や、曲線がある程度の傾きをもって単調増加していない場合は、その問題に何らかの検討すべき課題があることを示唆している。

3.4 節の最後で触れたように、表 5 の基礎統計量と図 6、表 8 を組み合わせると、各学部の実験者の学力を識別可能なテストになっていたかどうかを確認することができる。表

5 をみると、C 学部の合計得点は平均 55.7、標準偏差 14.4 で分布している。この分布を表 8 の学力レベルに照らし合わせると、C 学部の受験生は学力レベル 4~5 に属することがわかる。これに対応する図 6 の Lev4~Lev5 をみると、得点率はすべての大問について右上がりになっているので、大問 1~3 で構成されるテストは C 学部の受験者の学力を識別可能だったことがわかる。とくに、大問 2 は Lev4 と Lev5 間の得点率の差が大きく、他の 2 問よりも C 学部の受験者の学力の識別に寄与したことが推察される。

表 6 テストの信頼性

信頼性係数の推定値 (Cronbach の α)	0.733
-------------------------------------	-------

表 7 大問の特性

	得点率	識別力	削除時 α
大問 1	53.9	0.500	0.717
大問 2	32.4	0.595	0.604
大問 3	40.4	0.598	0.595

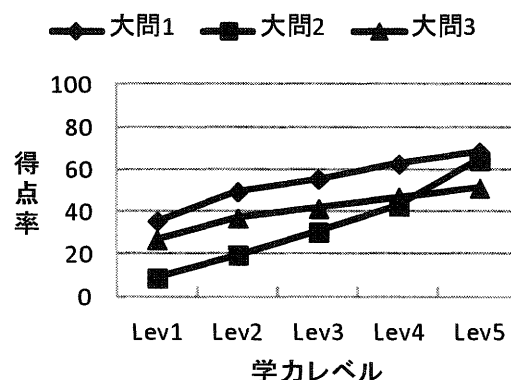


図 6 大問得点率分析図

表 8 大問得点率分析図の受験者群

学力レベル	受験者数	合計	
		平均	標準偏差
Lev1	204	24.3	8.8
Lev2	205	38.6	2.1
Lev3	213	45.9	2.0
Lev4	193	54.0	2.7
Lev5	185	65.4	4.8

3.6 相関係数と共分散比

表 9 に大問得点間の相関と共分散比を示し、表 10 にセンター得点と大問得点との相関を示す。表 9 と表 10 をみると、試験の設計目的に沿った試験結果が得られたかどうかを検証することができる。

大問得点間の相関係数により、その科目が測定する学力の構造をみることができる。表 9 をみると、大問 1 と大問 2 との相関係数の値は 0.441、大問 2 と大問 3 との相関係数は 0.560 であり、大問 3 は大問 1 よりも大問 2 に近い側面を測定しているといえる。また、大問 3 と大問 1 との相関係数は 0.443 であり、これらの大問は互いにある程度重なりながらも大問ごとに独自のものも測定していることがわかる。たとえば、英語のリーディングの大問どうしの相関は高く、ライティングの大問との相関はそれより低いなど、各大問の出題内容の整合性を確認することができる。

共分散比は、得点全体のばらつきのうち、各大問の得点のばらつきが占める割合を表す。共分散比の値が大きい大問ほど合計得点の得点差への寄与が大きい問題とみなされる。表 9 をみると、合計得点の得点差について、大問 1 より大問 2 と大問 3 の寄与が相対的に大きかったことがわかる。別の表現をすると、このテストでは大問 1 よりも大問 2 と大問 3 によって得点差がついたといえる。ところで、大問 1~3 の配点はそれぞれ 30, 30, 40 であり、作題者の意図としては大問 3 をより重視して

いることがわかる。この例では許容範囲と考えられるものの、配点比と共分散比が大きく異なる場合、現実の試験結果がテストの設計目的を反映していないことを意味する。

表 10 は、対応関係にある科目どうしのセンター試験と個別学力検査の相関関係をまとめたものである。たとえば、センター試験の「英語」と個別学力検査の「英語」が対応関係にある教科の例である。表 10 では、センター得点と個別学力検査の合計得点に強い相関が認められるので、両者の試験は比較的近い学力の側面を測定していたといえる。各教科の設計目的にもよるものの、個別学力検査はセンター試験と中程度から強めの相関があることが望ましいと考えられる。この例の場合、センター試験との相関から、統計的には適切な学力の側面を測定する個別学力検査であったといえる。

表 9 大問得点間の相関と共分散比

	大問 1	大問 2	大問 3	合計
大問 1	1			
大問 2	0.441	1		
大問 3	0.443	0.560	1	
合計	0.735	0.849	0.837	1
共分散比	0.252	0.389	0.359	1

表 10 センター得点と大問得点との相関

	大問 1	大問 2	大問 3	合計
センター	0.535	0.680	0.648	0.773

4 今後の課題

本稿では、個別学力検査における作題の PDCA サイクル構築に向け、入学センターで検討した個別学力検査の作題者へのフィードバック内容について分析例を提示した。今回の試みでは、大問得点のみを活用し、本学の個別学力検査における試験結果の一部を利用してフィードバック内容を第 3 節にリストア

アップした。今後、実際に個別学力検査の分析結果を作題者にフィードバックするとともに、その内容の改善についても検討していく予定である。

今後の課題の一つとして、フィードバック内容に合否の情報を含めるかどうかという問題がある。たとえば、学部別の得点分布を合格者・不合格者別に示せば、各大問の合否への影響を学部単位で把握することができる。一方、作題者にとっては情報過多になる可能性もあり、実際にフィードバックを実施した上で現場の意見を反映させることが望ましいと考えられる。

第二の課題として、表1の「物理」のように受験パターンが複数ある場合、代表的な受験パターンだけについて分析すると、どうしても統計情報をフィードバックできない大問が出てきてしまうという問題がある。一つの解決方法としては、「○」がついていない部分は欠測データとみなせるので、何らかの方法で大問得点の欠測値を埋めてから、すべての学部のデータを用いて分析することが考えられる。

第三の課題として、小問得点の活用があげられる。小問得点を利用することにより、さらにきめ細かな試験結果の分析が可能になる。ただし、第一の課題と同様に、作題者にとって情報は多ければよいというわけではなく、

フィードバック内容の取捨選択が重要な課題となる。

参考文献

- 林 篤裕 (2001). 「大学入試センターにおける統計作業」 『第1回S-PLUSユーザカンファレンスパワーポイント資料』 <<http://www.msi.co.jp/splus/support/salon/userconf/ronbun/1st/hayashij.pdf>> (2012年3月8日)
- 肥田野直 (編) (1972). 『心理学研究法第7巻テストI』 東京大学出版会.
- 池田 央 (1973). 『心理学研究法第8巻テストII』 東京大学出版会.
- 新潟大学 中期目標・中期計画一覧表 文部科学省 2010年4月23日 <http://www.mext.go.jp/component/a_menu/education/detail/_icsFiles/afieldfile/2010/04/23/1292754_14.pdf> (2012年3月2日)
- 適性試験委員会 (編) (2011). 『法科大学院統一適性試験テクニカル・レポート2009-2010』 商事法務.
- 適性試験管理委員会 2011年法科大学院全国統一適性試験実施報告書 日弁連法務研究財団 2011年12月28日 <<https://www.jlf.or.jp/jlsat/pdf/2011houkouku.pdf>> (2012年3月8日)