

測定・評価部門

教育心理学における測定・評価研究の一年

服 部 環

(宇都宮大学)

本稿は教育心理学における測定・評価研究について概観する。内容は、(1)教育心理測定に関する方法論的研究と応用的研究、(2)統計教育を含めた、新しい統計的データ解析法、(3)テスト問題の作成と実施方法を中心とする教育評価法、(4)測定尺度と心理検査の標準化である。対象としたのは日本教育心理学会、日本心理学会、日本教育工学会、日本行動計量学会の機関誌に掲載された論文および1991年度の学会総会で報告された論文、大学や研究所などが発行する研究紀要に発表された論文、そして、市販の書籍・心理検査である。いずれも原則として発表の期間を1990年7月から1991年6月までとした。

なお、総会の測定・評価部門で発表された論文をすべて取り上げるべきところであるが、他の学会で発表された論文も対象としたため、全部を盛り込むことはできなかった。

1. 教育心理測定

(1) 項目反応理論に関する方法論的研究

項目反応モデルは、能力値が大きい者ほど項目に対する正答期待確率が大きくなることを仮定するが、実際は能力値の高い者が誤答して、それよりも能力値の低い者が正答することは避けられない。しかし、頻繁にそうした不適合な現象が見られるときは、データのモデルへの当てはまりを疑わなくてはならない。例えば、孫・芝(1990)が日本人用に作成された語彙理解力検査を外国人に実施したところ、日本人と比べて外国人の反応パターンがモデルに適合しないことが示された。そこで、孫らは能力特性に変動を認める一般項目反応理論を提案し、語彙理解力検査の反応パターンへ適用している。その結果、2母数ロジスティック・モデルよりも適合度の高いことが示され、個人内変動を許した効果が顕著であった。提案されたモデルは個人特性に個人内変動を許した点に特徴があるが、そのことによって推定母数の数が増え、安定した推定値を得るためには、従来のモデルよりも数の多い項目を必要とする。もし実験に使用した程度の項目数を必要とするのであれば実用性に欠けてしまう。

なお、鈴木(1983)においても能力特性に変動を許す項目反応モデルが提案されている。特性値の分散をモデルの母数として導入している点は孫らのモデルと同様であるが、このモデルは特性値の分散が0となる場合にモデル曲線を定義できないので、孫らのモデルのほうが一般性は高い。

外れ値の多い反応パターンが出現したとき、その対処として孫らのように新しい測定モデルの提案もあるが、外れ値の存在を想定して個人の特性値を推定することができる。藤森(1991b)はそうした際の測定モデルと推定方法との組み合わせについて検討している。数値実験によっていずれのロジスティック・モデルにおいても、事前分布の設定が難しいものの、ベイズ推定法が適切であり、一方、外れ値の多い反応パターンに対して3母数ロジスティック・モデルが適切であると述べている。3母数ロジスティック・モデルが他のモデルより優れているのは、モデル化の違いから当然であるが、モデルが想定していない要因によっても外れ値が出てくることがあるので、それで十分ということにはならない。また、藤森が利用した方法以外にも頑健な推定方法があり、それを加えて検討することもできる。さらに問題を展開させて、標準誤差の推定方法についても検討することができよう。

多肢選択形式のテストは当て推量による正答を避けられない。藤森(1991c)は5肢選択式の項目をランダムに反復提示して、当て推量行動について検討している。実験の結果、能力水準の高い者ほど当て推量行動が少なく、能力水準よりも困難度の高い項目に対して当て推量の多いことが確認された。一方、当て推量による正答を除いた修正正答数と2母数ロジスティック・モデルによる能力推定値との相関は、3母数ロジスティック・モデルとの相関よりもやや大きかった。これは、この資料に関する限りわざわざ3母数ロジスティック・モデルを使う必要のないことを示している。おそらく、他の多肢選択式テスト・データにおいても、実用的には2母数ロジスティック・モデルがうまく機能するように思われる。

項目プールを作成するときは項目数が多くなるので、実際には項目をいくつかのテストに配分して、異なる受

験者集団に実施する。したがって、たとえ項目反応理論に基づいて項目母数の目盛りづけを行っても、異なるテストの項目母数の値をそのまま比べることはできない。一般にはテスト間に共通項目を入れ、それを要にして項目母数を共通尺度上へ等化する。この作業はコンピューター・プログラムを用いて自動的に行うことができる。しかし、そのプログラムを誰にでも利用できる環境にあるわけではないので、少々面倒であるが、テストごとに項目母数を推定した後に推定値を等化することがある。このような人のために、野口（1990）は異なる尺度上ですでに母数が推定されている2つのテストを共通の受験者に実施し、それを要として等化係数を推定する方法を提案している。以前にも野口は同様の問題を解決する方法を提案していたが、等化に必要な局外母数も推定していた。新たに提案した方法は尤度関数から局外母数を積分消去して構造母数のみを推定する、いわゆる周辺最尤推定法を利用している。シミュレーション実験によれば、受験者数が多い場合に適切な推定値を得られることが示された。しかし、計算コストが大きくなるためにシミュレーション実験が制約されて、実際の場面よりもバリエーションが少ないように思われる。他の方法との比較を含めて、提案した等化法の特徴についてさらに検討することが期待される。

自信度を補助情報とする項目反応モデルが繁樹・脇本（1991a）によって提案されている。特徴は、正誤反応と自信度をそれぞれ別個にロジスティック関数を用いてモデル化し、個人ごとに学力と自信度との間に2変量正規分布を仮定している点にある。実データの分析例によれば、項目の特性が明瞭になっている。自信度を2値的な方法以外で直接的に評定させることもできるが、項目困難度と自信度との関係をロジスティック関数を用いて視覚的に見やすく表現できる点が良いと思う。現在は項目ごとに学力と自信度との相関の大きさが異なるモデル、さらに、解答所要時間を組み込んだモデルを検討中ということなので（繁樹・脇本，1991b）、今後の成果を期待したい。

(2) 応用的研究

教材の構造は1次元的にはならないが、指導順序は1次元的に配列される。教師による質的な吟味によって教材を構造化するのはもちろん必要であるが、量的な指標によって教材の指導順序を検討する必要もある。こうした問題意識を持って、喜岡（1991）は小学校で学習する整数、小数、分数の四則演算に限定して、計算課題の困難度の推定を2母数ロジスティック・モデルによって試みている。これまでの指導要領では“分数×整数”が小学校5年生、“分数×分数”が小学校6年生に配当されて

いたが、平成4年度から施行される指導要領ではいずれも小学校6年生に配当されることになっている。当該の項目の困難度を調べてみると、新指導要領の配当の適切性がデータからも確認される。

同様に、藤森（1991a）は小学校3年生から5年生までの算数学力を測定する項目を、2母数ロジスティック・モデルによって尺度化している。能力値の分布を調べると、3年生と4年生との平均能力差よりも4年生と5年生との平均能力差のほうが大きかった。ところが、喜岡が作成した“能力測定値の累積分布”の50パーセンタイル値を見ると、むしろ逆の傾向が見られる。その理由は喜岡が計算問題だけを利用したことによるのか、標本の違いによるのか、あるいは他の要因によるのか、検討を要する。今後の課題は低学年用の問題作成である。また、紙筆式のテストを作成する上で、必要な項目を項目プールから自動的に選択し、反応パターンと能力測定値との対応表を作成するソフトウェアがあると便利である。

パソコン用の学習ソフトウェアが種々のメディアを通して発表されている。そのなかにはドリル形式でテストと学習を進めていくものがあり、受験者ごとに出题項目を替えている。しかし、提示項目の順序を無作為に決めているだけで、受験者の習熟度に応じて項目を選択しているわけではない。受験者の習熟度に応じた項目を提示するという観点に立てば、この領域にもテスト理論が貢献できる。許・繁樹（1990）が1つの例である。許らは教授内容をいくつかのステージに分け、さらに各ステージを下位クラスに細分化し、教授内容の階層構造化を図っている。そして、各ステージにおいて事前テストによって大まかな能力推定値を求めて、受験者が解答を進めるごとに能力値を推定し、各ステージの合格標準値を越えていれば次のステージへ進み、そうでなければ1つ前のステージへ戻って学習を行う、というテスト方式を提案している。構造化した教材に項目反応理論を適用した点と、逆戻りをする際にいずれのクラスへ戻って学習を行うべきか、クラスごとに優先度を定義した点が興味深い発想である。この方式では補助練習をすべきクラスを容易に決められるので、システム全体を未習得領域を見いだすための診断テストとして利用できる。ただ、この方式は用意した項目が多いときに威力を発揮するので、実験自体にやや物足りなさを感じる。

(3) 研究書

久しぶりにテスト理論を本格的に取り上げた研究書が出版された。日本語教育学会（1991）と竹谷（1991）である—さらに1冊あるがレビューは次年度の対象となる。前者は留学生のための日本語能力試験の基盤作りを進める過程で生まれた研究成果であり、そのなかで村上

(1991) が主に古典的テスト理論を取り上げている。日本語教育に従事する者を念頭に置いて書かれているため、文章による説明を優先させている。特徴は、議論の流れを断ち切らないよう考慮して、①本文の理解を深めるために必要な統計的基礎とエピソード的な題材を囲みの形で提供する、②詳しい証明は他書の該当箇所を参照できる、③一読目には飛ばして良い節と項を指示している、④節や項のタイトルを工夫している点にある。さらに、実際のテスト・データを素材として分析の流れが示されているので、心理検査の作成に関心を持つ者に具体的な手順がわかる説明となっている。信頼性と妥当性、また、テスト・データの分析法の理解を深めたい者には教わるところが多い。

古典的テスト理論と項目反応理論は、受験者の相対的な評価を目的として利用されることが多い。ところが、日常の授業では受験者数の少ないのが普通で、しかも、個人の到達度を質的に評価することが要請されるので、いつもそれら2つの理論が適当とは限らない。そこで、竹谷(1991)は個人の質的な評価に視点をおいて、分析手法を数理的に整理した。形成的テスト・到達度テストの分析、態度・心理の構造分析を中心として、S-P表分析、尺度解析法、項目関連構造分析法、意味構造分析法、ソシオメトリ分析法を説明している。手法の数理的な展開に終わらず、具体的な事例が付されているので、実用書としても利用できる。なお、本書が取り上げた“教育情報の構造分析法”がテスト理論の1つとして一般に認知されるには、もう少し時間が必要だと思われる。

2. 統計的データ解析法

(1) 共分散構造分析と因子分析

因子分析は探索的な分析手法として用いられることが多いが、ターゲット行列を特定して、それに近づけるように因子軸の回転を行う確認的な利用もされている。共分散構造分析は、やや正確さを欠くがこうした確認的あるいは検証的な因子分析とパス解析とを同時に扱うことのできる分析手法である。共分散構造分析モデルとしては、LISRELモデル、Bentler & Weeksモデル、COSANモデルが有名である。

豊田(1990)はこの3つのモデルの説明を通して、LISRELモデルとBentler & Weeksモデルが同等で、LISRELモデルがCOSANモデルに包含されることを証明した。さらに、3つのモデルの問題点を指摘し、①下位モデルの集合がLISRELモデルやBentler & Weeksモデルと同等で、②外生変数と内生変数、観測変数と潜在変数、構造変数と誤差変数、の区別を残して、③内生的潜在変数と外生的潜在変数の両方から同時に影

響を受ける観測変数を扱うことのできる、共分散構造の新しい表現法を提案している。ポイントはLISRELモデルの測定方程式を1つにまとめたところにある。そして、ITPA言語学習能力診断検査の下位尺度間の解析に、新しい表現法を見事に適用させた。

3つのモデルの他に共分散構造を表現する方法としてRAMモデル(McArdle & McDonald, 1984)がある。RAMモデルは推定すべき母数行列がLISRELモデルよりも少ない点に特徴があり、豊田がこのモデルも含めて共分散構造の表現法に考察を加えていれば、さらに新しい知見が得られたに違いない。今後の考察を期待する。また、豊田はLISRELモデルの問題点を指摘した上で新しい表現法を提案し、しかもLISRELプログラムを用いて計算している。LISRELプログラムを利用したときのコマンドが掲載されていれば、ユーザの参考になるだけでなく、新しい表現法が利用される機会も増えるであろう。

なお、LISRELプログラムは単体としてパソコンで利用できるが、大型汎用機用SPSSのオプション・プロダクトにもなっており、利用しやすくなっている。また、SASのリリース6.04以降を使えば共分散構造分析を行うことができる(豊田, 1991a)。共分散構造分析の説明と実際に利用した例が青木(1988)、清水(1989a, b)、柴田(1991)、土田・二宮・平野(1991)にもある。

プロクラステス回転法を用いて確認的な因子分析を行う場合、一般にはターゲット負荷を0あるいは1とする。しかし、南風原(1991b)によれば、それには属性と項目との関係の強さを一律に扱う点に問題がある。報告は、測定しようとしている属性について十分な知識を持っている者に属性と項目との適合性の判断を求め、適合すると判断された割合をターゲット負荷として用いる項目分類法を提案している。尺度の表面的妥当性と内的整合性を同時に高めるユニークな尺度構成法である。適用事例では、標準誤差の大きさを統制して標本数が決められており、標本数の決め方の例示ともなっている。

また、柏木(1991)は確認的な因子分析の重要性に触れ、プロクラステス回転に関するこれまでの研究成果を盛り込んだ、因子回転のプログラム・パッケージを提案している。それを用いると、ゼロマックス回転から不完全直交あるいは斜交プロクラステス因子回転までが可能である。利用しやすい形態でユーザに提供されることを希望する。

近年、多相データの解析手法に関する研究が進んでいる(吉澤, 1991)。吉澤によれば、①要因実験の多元配置データ、②多重分割表、③多重配列0-1データ、④官能評価データ、⑤類似度データ、⑥多変量縦断データにおい

で多相データ解析が可能である。多相データとしてはSD法データが身近な例であるが、実際の解析では3相データを2相データへ圧縮して因子分析を行うことが多い。本来ならば3相データのままで解析を進めたい。1つの例として、3相データの因子分析を行う最も単純なモデルにPAFAFACモデルがある。パソコン用の計算プログラムは服部(1991b)にある。探索的な解析には向かないが、モデルに課された制約を理解しておけば3相データの解析に利用できよう。

因子分析について、行動計量学誌に理論的な観点から特集が組まれている(繁樹・宮埜, 1990)。狩野(1990)は統計的推測の発展を概観し、標本数が十分に大きいときは最尤法、そうでないときは単純最小自乗法や反復推定を必要としない推定法を推奨している。小笠原(1990)は共分散構造分析について、一般的なモデルが扱わない領域に重点を置いて説明している。共分散構造分析を概観している論文ではないが、因子分析と回帰分析とを統合したモデルを通して、因子分析の理論的な展開を知ることができる。また、3相データの因子分析・主成分分析の発展を村上(1990)が述べている。繁樹(1990)はカテゴリカル・データの因子分析のモデル化の手順と母数推定法を概観し、今後の課題を論じている。これらは刺激的な論文であるが、理論に重点をおいたため、論文を読んだだけでは共分散構造分析、3相データの因子分析、カテゴリカル・データの因子分析などをすぐに利用するのは困難である。特集で論じられた手法が一般に利用されるためには、使い勝手の良いソフトウェアと実例を豊富に盛り込んだ概説書が必須である。

(2) 新しい統計的解析法

豊田(1991a)は2次元配列の度数表を解析するために、“行側の刺激と列側の刺激が共通の潜在特性上で定義され、刺激間の距離が短くなるほどそれが同一の刺激とみなされる確率が高くなる”，という項目特性曲線に基づいた数理モデルを提案している。項目特性曲線は位置母数、識別力母数、反応確率の上限を与える強度母数によって定義されるので、数量化Ⅲ類よりも詳しい解析が可能である。しかも、比較的少数のデータからも安定した解が得られ、モルスの混同行列の適用に見られる通り、多次元尺度解析法と同様の解析も可能である。確かにこのモデルは数量化Ⅲ類よりも優れている点が多いが、データごとに分析モデルをユーザが特定してから解析を進める必要があるため、数量化Ⅲ類のように手軽に利用するのは容易ではない。

数量化Ⅲ類と同様の解析法として双対尺度法がある。Nishisato(1991)は双対尺度法の可能性と限界について、①双対性、②行・列空間の違い、③データ・タイプ

と感度、④多次元空間のゆがみ、⑤拡張された次元と数量化の行き過ぎ、⑥等価分解の原理、⑦多元的多次元解析の可能性、の観点から述べている。限界として指摘されたなかで、特に心得ておくべきことは、行要素と列要素とを1つの図へ書き入れる際に、双対尺度流の重みとコレスポンデンス分析流の重みでは布置の解釈の仕方が変わること、欠損値の影響を受ける程度がデータ・タイプによって異なること、連続変数のカテゴリー化も含めて、カテゴリーの合併によって1つの因子が複数の因子になってしまうこと、などである。

量的変数を合成する方法として主成分分析がある。また、量的変数と質的変数の混在を許して合成変数を構成する方法にPRINQUALがある。大津(1990)が提案したOSMOD2もそれと同様に、質的変数として順序変数と名義変数、さらに量的変数を混在させて主成分分析を行うことができる。数量化Ⅲ類や双対尺度法では、反応頻度の小さいカテゴリーに対する重みが大きくなりすぎてしまう問題点が指摘されるが、OSMOD2はそれを改善し、さらに数値最適化を向上させている。今後、実データを用いた報告を望みたい。

椎名・松浦(1991)は、椎名自身の提案による集合論的潜在特性理論を一種の潜在クラス・モデルへと発展させている。モデルでは潜在クラスごとに項目正答確率が集合論的潜在特性理論に基づいて定義され、各クラスの事前確率、各クラスごとの問題正答確率、問題に正答するために必要な下位スキルの通過確率が算出される。報告ではモデルが割り算問題の解答パターンの分析に適用され、割り算問題に正答するための学習の道筋—発達のモデル—が吟味された。

パス解析を利用したあと、結果を見やすくするために説明変数から目的変数へ矢印を引くことがある。通常、両変数ともそれほど多くはないので、矢印が交差しないように線を引くことは難しくない。ところが、教材構造分析のように変数の数が増えると、矢印と矢印とが交差して、構造全体を見やすい図へまとめるのが楽ではない。それを解消するために、赤堀(1991a)は矢印の交差をなるべく少なくして線引きを可能にするアルゴリズムを提案し、学習指導要領に記載されている順序で作成した学習目標の系統図よりも見やすい系統図を作成した。他の手法との比較検討があればさらに興味深いものとなった。

さらに、赤堀(1991b)は1対比較データの順序尺度構成法のアルゴリズムの妥当性を、完備な応答行列と不完備な応答行列のいずれにおいても最適解を導くことができることにより検証している。これまでも1対比較データの尺度構成法が提案されているが、この方法の特

徴はノード間に逆順序を認める点にある。赤堀が指摘しているように上位概念を学んでから下位概念を学ぶこともあるので、本法の有用性は高く、新しい方法として評価される。

(3) 統計教育

南風原 (1991a) は6年前に行われたシンポジウムと重複する部分があると前置きして、“有意性検定からの脱却は可能か”と題する小講演を行った。教育心理学研究誌では、頻繁に有意差検定の結果が掲載されるが、そこには6年前のシンポジウムの効果が見られないと嘆いている。例えば、標本数を増やせば帰無仮説を棄却できてしまうのはよく知られた事実であるが、2分法的な検定を止めて効果の大きさの区間推定を行おうとしない。また、先行研究と同程度の効果の大きさが認められても、標本数が少ないために検定結果が異なってしまう、研究の積み上げが難しい。逆に有意な結果が得られたとしても効果の大きさが異なる場合もある。いずれにしても2分法的な検定をしている限り研究成果の蓄積は困難で、検定の枠内での改善策として、帰無仮説を一定の区間を持つ仮説とすること、それと同時に検定力を特定の値に設定して標本数を決めることを提言している。さらに、検定に代わる方法として、効果の大きさの指標の区間推定を行うことを強調している。これは2次分析やメタ分析のために必要な情報を蓄積するという、池田 (1990) の提言とも共通している。最後に今後の課題が述べられたが、まず誰かが教育心理学研究誌に2分法的な検定に頼らない論文を載せる必要があらう。

藤森・重松 (1990) は平均値の差の検定に焦点を当て、学習支援システムを作成している。この利用者は、すでに統計学を履修して基本的な用語についてある程度の知識を有している者である。特徴は、①検定法を適切に選択させる、②特にパソコンの知識を必要としない、③資料と選択した検定法との矛盾に対処している、④実際の計算もできる、点である。藤森らはシステムを利用した者が適切な検定法を選択できることを実験を通して確かめている。実際には、教育心理統計を受講した者でも、手元の資料に対して適切な検定法を選べないために、折角の統計パッケージを利用できないことがあり、補助教材としてこの援助システムがさらに充実していくことを期待したい。

ここで書籍についても触れておく。まず、倉知・山上 (1991) は教育心理統計の要点をまとめた教科書であるが、パソコン版 SPSS (山本・吉村・竹村, 1991) について入門的な説明があるのが特徴である。また、東京大学教養部統計学教室 (1991) は幅広く話題を取り上げているが、初学者には半期でこなすのは難しい。シリーズの

一巻として人文学生向けの教科書が予定されているので、その完成を待ちたい。林 (1991) は調査法にも触れているのが特徴で、例題が豊富である。また、飯島 (1990) は SAS を用いた著者自身の教育実践を踏まえてデータ解析の入門を行っており、工夫が見られる。また、菅 (1990) は推測統計から多変量解析までのパソコン用プログラムのソース・リストを掲載しており、プログラム言語も学ぶことができる。さらに、以上の書籍とは趣を異にしたものに市川 (1991) がある。心理学の各領域で行われている測定を述べており、測定という観点から眺めた心理学の入門書と言える。

3. 教育評価法

(1) 適応形テスト方式

Rasch モデルは項目の弁別力を等値とする、厳しい制約条件を課したモデルである。したがって、データへの適合度は他のモデルを超える可能性は小さい。一方、テスト得点が能力値の十分統計量になるという都合の良い性質を持ち、項目困難度の等化も比較的容易である。服部 (1990) はそうした利点を生かして、語彙理解力を測定する検査項目を作成している。さらに、パソコンを用いて受験者の能力水準に応じた項目を提示して、語彙理解力を測定する適応形テストも実施している。実験によれば、比較的信頼性の高いことが認められているが、問題点はデータとモデルとの適合度である。

井上・舛田 (1991) は検査項目を減らして実施時間を短縮するために、2母数正規累積曲線モデルに基づいて項目母数を尺度化し、パソコンを用いたベイズ方式による適応形性格検査を試みている。この方式は特性値の事後分散を最小とする項目を項目プールから1項目ずつ提示していくもので、任意の推定精度が得られたところで検査を終了できる。実験によれば、終了基準を満たした時点での推定値と全項目による素点との相関が十分に大きく、推定精度を落とさずに実施項目数を削減できることが示された。しかし、テスト情報量の大きい特性値付近の被験者が多かった可能性もあるので、さらに終了基準の適否を吟味する必要がある。それには特性値の大きさを变化させたシミュレーション実験が1つの検討方法とならう。

(2) 大学入試におけるテスト作成

鈴木・山田・池田・赤木 (1991) と石塚・前川 (1991) は共通一次試験で実施された国語と英語の問題を記述式に書き換えて、学力水準が等質と考えられる2つの集団の一方へ多肢選択式問題、他方へ記述式問題を実施した。その結果、テスト全体ではいずれの科目とも記述式のほうが困難度は高くなるが、テスト特性曲線を用いた分析

によれば識別性は類似していることがわかった。しかし、国語の項目によっては記述式に書き換えることでやや識別力の低下が見られ、英語の発音問題では、信頼性は多肢選択式のほうが高いものの、多肢選択式が再認力を測定しているのに対し、記述式では綴りの再生力も要求することになり、異質な学力を測定していた可能性があった。おおむね多肢選択式と記述式が等しい信頼性を持つことが指示されたが、英語に関しては多肢選択式問題では文章表現力や聴取力などを直接測定できないので、測定内容と信頼性とを分けて考えておくべきである。

一方、数学は多肢選択式といっても数字を答えさせているので、再認力を測定する問題とはなっていない。そこで豊田・山村(1991)は設問の細分化の影響を検討している。その結果、記述式のほうが困難度が高くなること、困難度の統制はマーク式のほうが容易であること、問題形式(マーク式対記述式)と設問過程(細分化するか否か)との交互作用が有意でないことがわかった。

以前、共通一次試験の理科で得点調整が行われたことがある。原因は科目間の難易度に関がかったからで、事前に平均を推測できればこのような事態は起こらない。山田・岩坪(1991a,b)は困難度の推定方式を開発するために、理科と社会の問題の定性的な側面に着目して正答率の予測を試みている。推定式は昭和62年度から3年間の資料を用いて導出され、平成2年度の資料で検証された。それによると、枝問正答率の重相関係数が比較的大きな科目もあったが、推定平均点は実際よりもやや低めに出ており、十分な予測力とはいえなかった。しかし、事前に困難度を推定するという興味深い着想であった。

ところで、それとは逆に、平均点が同じでも得点調整を必要とすることがある。つまり、特定の科目を学力水準の高い集団が受験した場合で、他の科目との平均点が同じならばその科目を選択した者が不利になるからである。特にこの問題についての説明は発表されていないが一例えば受験案内には科目間平均差の大きい場合についてだけ説明がある一大学入試センターの考えを聞いたものである。

(3) 大学入試制度

柳井・前川・豊田・鈴木・仙崎・赤木・中島(1991)と柳井・前川・豊田・鈴木・前田(1991)は共通一次試験の志願者が1名以上いた高等学校を被験校として進学指導の実態調査を行っている。前田・豊田・柳井(1991)は、調査で得られた大学入試制度の改革案に関する自由記述による回答をカテゴリー分類して、単純な制度対入試の多様化、直接的改善対社会状況からの間接的改善、現状の肯定対ランクづけに代表される現状の否定、という3つの軸を構成している。それらは自由記述を数量的

に分析する手順の良い例示でもあるが、標本の無作為性と本音をどこまで引き出せたかが課題と言える。

さらに、豊田・柳井・前田(1991)は意見の多かった推薦入試、入試科目数、一芸入試、調査書重視、受験機会の増加について、賛成意見と反対意見を手際よくまとめている。いずれの報告も入試業務に携わる者には是非フィードバックして欲しい内容である。発表の場では大学入試センターの姿勢を問う意見も出され、入試の話題を中心として活発な議論が行われた。入試の問題を共通の問題意識とするためにも、有意義な討論であった。

(4) 評価と評価用語の認知

会社訪問を受け付けて貰えなかった学生が、偽って有名大学の名前を出したところ即座に面接の日取りが決まったという話を聞いた。本来、大学入試の機能は各大学において選抜機能を果たせば良いのであるが、実情はこのように企業における選抜機能までも果たしている。こうした枠組みは容易に変えられないが、藤森・繁樹(1990)が指摘するように、在学中の成績が社会的に評価されるようになれば、いくらかは良い方向へ向かうかもしれない。藤森らは評価に関する改革の必要性を述べたあと、大学生を対象とした授業に関する調査結果を考察している。調査は講義スタイル、評価方法とその効果、成績評価において重視すべき要因、評価法の改革に対する賛否、学生による教師の評価を導入するとき考慮すべき要因などに亘っている。講義に不満を持っている学生が多いためか、“学生による授業評価”が一本邦では慶応義塾大学湘南藤沢キャンパス、筑波大学、文教大学情報学部が実施(検討)していると聞く、本邦の学生にも支持されていて興味深い。授業評価は教師にとってあまり気持ちの良いものではないが、学生への対処を知る上で参考になろう。進学率が数パーセントの頃は“わからなければ学生が悪い”という風潮であったようだが、今は教育機関として考え方を変える時期に来ていると思う。

織田(1991)は授業評価を教師だけでなく、学生の受講意識を変える有効な方法と捉え、教育学部における20の授業に対して学生による授業評価を行った。質問紙の回答分析の結果、学生は授業充実、学生参加、話術、授業内容精選の観点から授業を評価していることがわかった。授業評価の結果が各教官へフィードバックされたかどうかについては触れられていないが、授業評価の実践活動としても注目される報告である。

藤岡(1991)は教育評価と学習指導用語のイメージを小学校教諭を対象として調べている。10個の概念を取り上げているので、全体をまとめるのは難しい。全般的にほぼ10年前に福岡教育大の井上が実施した調査結果よりも評定値が中立化していた。

(5) 学業と動機づけ

下羽・三浦(1991)は学習への興味関心や学習習慣などの学習関連人格特性を想定した測定尺度を構成し、教科全般の“学業不振”との関連を示唆している。しかし、算数については尺度と標準学力検査の成績との関連は見られず、作成した尺度では捉えきれない知的能力の関与が予想された。分析では学業成績の5段階評定値から機械的に学力段階を分類しているが、“学業不振”の状態を個々に把握しておけば、作成した尺度を指導に生かすことができると思われる。

坂本(1991)は小学校5・6年生の学習スタイルを8通りに分類して、学習成績と学習スタイルとの関連を検討している。坂本のアイデアは東学大の岡本が作成した学習スタイルの3つの分類軸を両極性へ変容した点にある。しかし、測定尺度そのものに大きな修正はなく、単極性の学習スタイル分類尺度によって、両極的な学習スタイルを測定すべきかどうか、尺度の使い方に疑問が残った。今後、8通りの類型に対応した学習指導は困難と思われるが、考察が具体的な指導へつながればこの研究が生きてこよう。

杉村・伊谷(1991)は中学3年生の生活実態と学業成績との関係を検討している。生活実態に性差が認められているので、学業成績との関連を吟味する場合にも性を要因の1つとして含めたほうが良い。研究の目的が実態調査そのものであるが、指導との関連を考えておかなければ結果が生かされないように思われる。なお、報告のなかで“知能教科”、“技能教科”という言い方をしているが、適切な用語とは思われない。

鹿毛(1990)は自己評価対他者評価、相対評価対個人内評価を取り上げ、大学生を被験者として内発的動機づけへ与える効果を検討している。実験課題が知能検査問題であり、課題自体にはやや生態学的な妥当性を欠くが、実験状況を自然な状態へ維持するように努めている。実験の結果、他者評価が自己評価よりも内発的動機づけを低下させ、逆に不安を高めることが示された。一方、相対評価は個人内評価よりも内発的動機づけに基づく行動の一部を喚起していたが、他の内発的動機づけ、成績目標の設定、認知的干渉、有能感では相違が見つからなかった。相対評価と個人内評価との間に相違が小さいのは、鹿毛が考察しているように双方とも外的な基準を設定していないことが理由と思われる。実験設定に工夫が見られるが、被験者には課題の失敗だけを経験させているので、成功経験と内発的動機づけとの関係も検討されることが課題となろう。測定・評価領域において注目されるべき研究の1つである。

(6) 知能とその他

藤田・前川・小林・大六(1991)は、WAIS-Rの標準化データの年齢群別の相関係数行列から平均相関係数行列を求めて下位検査間の因子構造を探っている。2因子を抽出したところで言語性因子と知覚的体制化因子が抽出され、さらに3因子を抽出すると米国版の結果と同様に記憶・被転動性からの解放因子が抽出された。また、前川・藤田・小林・大六(1991)は9つの年齢群を青年期、成人前期、成人後期、老年期の4期に分けて、平均相関係数行列から因子抽出を行い、加齢に従って一般知能の分散が大きくなることを示した。また、言語性因子と動作性因子とを確認するために2因子を抽出すると、老年期を除いてほぼそれに対応する因子を抽出できたと報告している。一方、老年期では2因子間の相関が大きく、2因子構造が適切でないと述べている。

また、梅永・前川(1991)はWAIS-Rを知能障害児・者と境界線児・者に実施して因子構造を分析している。素点による分析と評価点による分析が行われ、言語理解、知覚的体制化、被転導性の3因子が抽出された。

このように、WAIS-Rの因子構造をめぐる日本語版製作者が中心となって上記の結果を報告している。知能の因子構造を探ろうとする意欲的な研究であるが、因子構造の特定に向きない探索的方法を用いているため、いずれの報告も検証的な根拠が薄く、因子数と因子構造がはっきりしていない。例えば、因子数を特定するなら最尤因子分析を利用することができる一モデル選択にはAIC以外の統計量も利用できる。さらに因子構造をモデル化するならば共分散構造分析を用いることができる一因子分析ではターゲット負荷を0としてプロクラステス回転を行っても推定値が0になるとは限らないが、共分散構造分析を用いれば0のままで因子構造を特定できる。

1930年代の終わりに優秀知能児の研究が行われ、森(1991)はその研究で選抜された児童の追跡調査を行ったと述べている。長期に亘る研究で、貴重な資料が得られたと推測されるが、本文では追跡調査を行ったという記述に終始し、その結果は口頭で簡単な報告が行われたに過ぎない。興味深い調査であるだけに、少しでも結果を報告して欲しかった。

山口(1991)は集団主義的傾向と性格特性との関連を検討している。18歳から78歳までを対象としたため、以前に行われた青年層だけからの結果と一致しない点がある。年齢を統制する統計的な手法で分析を行えば、別の知見が得られたようにも思われる。

4. 測定尺度と心理検査の標準化

(1) 知 能

鉄道運転士には事故防止のための適性試験が課され、バッテリーには一般的な知的機能の水準を調べる検査が含まれている。小笠原（1991）は鉄道運転士用に作成された集団式成人用知能検査の開発について報告している。検査は、①集中力・持続力、知覚の早さ・正確さ、②推理力・空間における関係判断を測定する2つの領域に分かれ、小笠原自身が開発した分析モデルを用いた解析によれば、単純な作業課題には知的機能の発現を助ける持久力が関連し、それが20歳代よりも中高年のほうで有利に働くことが示唆された。報告は主に検査得点の標準化手続を説明したものであるが、分析モデルの適用例を示したものとも言える。発達データの平均と分散共分散を説明する興味深いモデルだけに、他のデータにも適用されることを期待したい。また、鉄道運転士に対する指導管理者の評価と適性検査との関連が芳賀（1991）によって報告されている。相関分析によれば特に強い関連は認められていないが、今後も研究を継続することを望みたい。

日本語版 WAIS-R が改訂された（品川・小林・藤田・前川，1991）。検査は11の下位検査から構成され、16歳から74歳までの言語性知能と動作性知能を測定できる。手引き書には標準化の過程と検査の統計的な特徴が説明されているので、信頼性はそこから知ることができる。しかし、妥当性に関しては原版の説明があるのみで、日本語版の説明はなく、利用者に検証が期待されている。なお、下位検査ごとに評価点の平均と標準偏差が掲載されているが、むしろ、素点の平均と標準偏差のほうが読者の参考となる。

(2) 人 格

ロールシャッハ・テストのパソコンを用いた自動診断システムが発表された（村上・村上，1991a）。それは記号を入力するだけで各種の指標と自動解釈レポートを作成する。著者らによるとレポートの精密さはおそらく世界最良で、熟達したテスターを遙かに凌いでいる。とかく主観的になりがちな投影法の解釈法を標準化した、野心的なシステムである。

MMPI は同一原版に対して本邦ではいくつかの翻訳版が出ている。小口・高見堂（1991）は市販されている日本版と研究用に作成された同志社版とを、翻訳内容と項目反応から比較検討している。その結果、翻訳版相互の間には越え難い隔りがあり、特に前者には翻訳上の問題が多く、それを統一版とはできないと述べている。

一方、村上・村上（1991b）はMMPIの原版を改めて翻訳し（MMPI-1）、信頼性と妥当性を落とさずに質問項目数の削減を試み、最終的に250項目によって13尺度を構成し、さらに、250項目の因子分析を通して9尺度を構成

している。システムの有効性は大鬱病患者の診断を通して、MMPI-1との比較検討によって吟味されている（村上・村上，1991c）。それらは日本におけるMMPIの統一版の可能性を示唆する報告でもある。

米国のギルフォードらが作成した4種の性格検査から、12の特性尺度を選んで編集した試筆式の検査が作成されている。その検査がパソコンを用いて個人別に実施できるようになった（本明・久米・織田・応用教育研究所，1991）。実施後ただちに結果をフィードバックできる点に特徴があるが、個人の特性水準に合わせて実施項目を選択していない。使い勝手の評価は今後に待ちたい。

(3) そ の 他

尺度作成に関連して、服部（1991a）はもっぱら内部一貫性に着目して項目選択技法を吟味している。項目反応理論を用いてターゲットのテスト情報量を描き、その形状に実際のテスト情報量を近づけるような項目選択技法があれば便利であるが、その方法がそこまですぐに結びつくとは考えにくい。

米国で作成された承認欲求測定尺度の妥当性の問題が指摘されるようになり、その後、新たに尺度が作られた。植田・吉森（1991）はそれを日本の大学生へ実施して、日本語版でも信頼性の高いことを確認した。しかし、因子的妥当性に関しては原版とは異なる因子構造が得られており、その原因が抽出法と回転法の相違によるものなのか、もともと因子構造に相違があるのかがはっきりしない。相違の理由がわかるような解析手順が良かったと思う。

中学生と高校生の適応性を診断する検査が標準化されている。検査の特徴は、環境と自分自身を認知する過程と、それに基づいて反応する過程とに分けて適応性を診断する点にある。中原・高森（1991）は教示文を児童用に改め、標準化の準備を試みているが、小学生に120項目を課するのはやや負担が大きく、信頼性と妥当性を大きく低下させない範囲で項目を削る必要があろう。また、一律に得点で診断を下すのは危険なので、診断結果を外的な基準と照合しておく必要がある。

進路適性検査の判定基準が改訂された（相談と検査の研究会，1991）。検査は就職・進学両分野の進路を同時に判定して、結果は教師と生徒に返却されるので、進路指導の教材として利用することができよう。

5. おわりに

最後に、ここで取り上げられなかった報告も含めて、感想を述べたい。ただし、検定結果に対して過剰な解釈が見られた報告があったが、その点については触れる必要はないだろう。

まず、教育心理測定に関しては今年も項目反応理論に関する理論的・応用的研究が発表された。すでにこの理論がテスト作成の標準的な測定理論になりつつある学会もあり、実践データの蓄積が進んでいる。従来からのテスト理論が不要になるとは考えられないが、測定・評価部門に関心の浅い者でも、項目反応理論を利用できる場面は多いはずで、さらに注目されて良いと思う。

新しいデータ解析手法が提案されたとき、論文を通して手法の優れていることが理解される。今年度もそのような論文が発表されたが、コンピュータ・プログラムの作成がユーザに委ねられているため、簡単には提案された手法を利用できない。そこで、コンピュータ・プログラムがフォーラム・レポートに登録されて、一般ユーザへ提供されることを望みたい。

例年通り、因子分析が各領域で頻繁に利用された。利用形態としては探索的な利用がほとんどで、研究の流れから検証的な解析が適している場合にも探索的な解析が行われている。例えば、WAIS-Rの因子構造に関する研究やパーソナリティの5因子説に関する研究(辻, 1991)では一研究の目的通り分析結果はいずれも見事な単純因子構造が得られたが、仮説とした因子構造と共分散行列との適合性を検証していく方法もあったと思う。検証的な分析手順を通してモデルの構築が進むことを期待したい。

また、統計的検定に関しての小講演で有意差検定を行う上での問題点が指摘されたが、改善策の1つである、“専門家自身がやってみせる”ことを希望する。これが大事なのは、統計的検定を離れてしまうが、因子分析でプロマックス回転法が頻繁に利用されるようになった経緯を考えてみれば容易に推察されよう。

実態調査が必要なときもあると思うが、報告のなかには実態調査に終始しているものが見られた。調査項目を増やして実態を探り、あるいは特性値の大きさから被験者を細分化して特性値間の関連を見ても、学習・生徒指導にどのように生かすかを考えて調査を進めなくては、研究結果が生かされないように思う。特に測定・評価領域の調査では指導へのつながりを念頭においた研究が望まれよう。

引用文献

(注) 教心総会は日本教育心理学会第33回総会発表論文集、日心大会は日本心理学会第55回大会発表論文集を指す。

青木幸弘 1988 潜在変数を持つ構造方程式モデル — LISREL アプローチの概要 情報科学研究, 3, 49-62.

赤堀侃司 1991a 学習目標系統図の資格的表示における交差数減少の一方方法 日本教育工学雑誌, 15, 1, 25-34.

赤堀侃司 1991b 1対比較法の逐次計算法と妥当性の検討 教心総会, 773-774.

藤森 進 1991a 小学校3年生から5年生の算数学力尺度の作成 心研, 62, 2, 82-87.

藤森 進 1991b 被験者の反応パターンと能力推定値 岡山大学教育学部研究集録, 86, 45-51.

藤森 進 1991c 多肢選択テストにおける「あて推量」の発生に関する一研究 岡山大学教育学部研究集録, 87, 41-48.

藤森 進・繁樹算男 1990 大学における評価と授業法に対する学生の意識調査 日本教育工学雑誌, 14, 2, 97-103.

藤森 進・重松 宏 1990 統計的検定法の選択のための援助システムの効果 日本教育工学雑誌, 14, 3, 121-129.

藤岡秀樹 1991 教育評価の研究(3)ー小学校教師の教育評価や学習指導に関する用語の認知についてー, 教心総会, 755-756.

藤田和弘・前川久男・小林重雄・大六一志 1991 日本版 WAIS-R の因子構造に関する研究(1)全被験者群による探索的因子分析ー, 日心大会, 851.

南風原朝和 1991a 有意性検定からの脱却は可能か, 教心総会, L17-L18.

南風原朝和 1991b 因子分析を用いた項目分類における主観的判断データの利用, 教心総会, 779-780.

芳賀 繁 1991 鉄道運転士のための適性検査の妥当性について, 日心大会, 857.

服部 環 1990 個人差に応じたテスト方式による語彙理解力の測定, 教心研, 38, 445-454.

服部 環 1991a テストの内部一貫性を大きくするための項目選択技法, 教心研, 39, 195-203.

服部 環 1991b HPARAFAC. EXE: PARAFAC モデルを解くパソコン用プログラム 教育心理学フォーラム・レポート, FR-91-002.

林知己夫(編) 1991 統計学の基本 朝倉書店

市川伸一(編著) 1991 心理測定法への招待 サイエンス社

飯島正樹 1990 SAS による大学生のための実用統計学入門 山海堂

池田 央 1990 行動計量学の資産を作ろう 行動計量学会報, 50, 1.

井上俊哉・榊田博之 1991 ベイズ方式による適応形性格検査の試み 教心総会, 791-792.

- 石塚智一・前川眞一 1991 英語の試験問題の出題形式に関する比較研究 大学入試センター研究紀要, 20, 47-72.
- 鹿毛雅治 1990 内発的動機づけに及ぼす評価主体の評価基準の効果, 教心研, 38, 428-437.
- 柏木繁男 1991 因子回転のプログラムパッケージについて, 日心大会, 638.
- 管 民郎 1990 入門パソコン統計処理 技術評論社
- 狩野 裕 1990 因子分析における統計的推測: 最近の発展 行動計量学, 18, 1, 3-12.
- 喜岡恵子 1991 項目反応理論に基づく計算課題の難易度の推定, 教心研, 39, 204-213.
- 倉知佐一・山上 暁 (編) 1991 要説心理統計法 北大路書房
- 許 紅・繁樹算男 1990 項目反応理論と授業内容の階層構造的表現による問題項目の提示順序の最適化 日本教育工学雑誌, 14, 2, 73-80.
- McArdle, J. J., & McDonald, R. P. Some algebraic properties of the Reticular Action Model for moment structures. *British Journal of Mathematical and Statistical Psychology*, 37, 234-251.
- 前田忠彦・豊田秀樹・柳井晴夫 1991 高校教育現場からの大学入試改善に関する意見の研究, 教心総会, 751-752.
- 前川久男・藤田和弘・小林重雄・大六一志 1991 日本版 WAIS-R の因子構造の変化, 日心大会, 851.
- 森 重敏 1991 優秀児の特質に関する基礎研究 (第28 報告) -大阪市抜群優秀知能児の追跡-, 教心総会, 769-770.
- 本明 寛・久米 稔・織田正美・応用教育研究所性格検査研究部 1991 本明・ギルフォード性格検査 ソフトウェアマニュアル 図書文化社
- 村上 隆 1990 多群・多集合・多相データの主成分分析について 行動計量学, 18, 1, 28-40.
- 村上 隆 1991 第2章「良いテストはどのような性質をもつか」, 第5章「大規模なテストデータの分析の一例」, 第8章-1「形成的評価の理論」 日本語教育学会 (編) 日本語テストハンドブック 大修館書店
- 村上宣寛・村上千恵子 1991a ロールシャッハ・テスト -自動診断システムへの招待- 日本文化科学社
- 村上宣寛・村上千恵子 1991b MMPI 自動診断システム(5) -MINI システムの作成- 日心大会, 846.
- 村上千恵子・村上宣寛 1991c MMPI 自動診断システム(6) -事例: MINI, MMPI-1 による大鬱病の心理学的評価-, 日心大会, 847.
- 中原弘之・高森 寿 1991 児童用適応性診断検査の作成について 茨城大学教育学部紀要 (教育科学), 40, 327-343.
- 日本語教育学会 (編) 1991 日本語テストハンドブック 大修館書店
- Nishisato, S. (西里静彦) 1991 Possibilities and limitations of Dual Scaling, 日心大会, S64.
- 野口裕之 1990 共通被験者デザインにおける等化係数の周辺最尤法による推定 名古屋大学教育学部紀要, 37, 191-198.
- 織田揮準 1991 「学生による授業評価」にみられる教授活動 1991年教育工学関連学協会連合第3回全国大会講演論文集, 317-318.
- 小笠原春彦 1990 共分散潜在構造モデル 行動計量学, 18, 1, 13-27.
- 小笠原春彦 1991 成人用知能検査 J-1001の開発, 教心総会, 757-758.
- 小口 徹・高見堂正彦 1991 解答傾向からみた MMPI 翻訳差の検討, 日心大会, 636.
- 大津起夫 1990 尺度混在の多変量データのための非線形主成分分析: OSMOD 2 HBSR-M (S)-13 北海道大学文学部行動科学科
- 坂本徳弥 1991 学習スタイルと学業成績の関係についての一考察, 教心総会, 747-748.
- 柴田 満 1991 親子の性格特性間の因果分析 -性格心理学における共分散構造分析の応用 日本 SAS ユーザー会論文集, 91-96.
- 繁樹算男 1990 カテゴリカルデータの因子分析 行動計量学, 18, 1, 41-51.
- 繁樹算男・宮埜寿夫 1990 特集「因子分析」について 行動計量学, 18, 1, 1-2.
- 繁樹算男・脇本あゆみ 1991a 学力と自信度の同時推定について 教心総会, 777-778.
- 繁樹算男・脇本あゆみ 1991b 所要時間及び正答確率を同時に説明するモデル 1991年教育工学関連学協会連合第3回全国大会講演論文集, 397-398.
- 椎名乾平・松浦三郎 1991 潜在確率を用いた割り算誤答分析, 教心総会, 781-782.
- 清水和秋 1989a 検証的因子分析, LISREL そして RAM の概要 関西大学社会学部紀要, 20, 2, 61-86.
- 清水和秋 1989b 中学生を対象とした進路不決断尺度の因子的不変性について -COSAN を使用して- 関西大学社会学部紀要, 21, 1, 143-176.
- 下羽美枝子・三浦香苗 1991 学習に関連する人格特性尺度の作成 -児童の学業成績・標準学力テスト結果と関連させて- 教心総会, 745-746.

教育心理学年報 第31集

- 品川不二郎・小林重雄・藤田和弘・前川久男 1991 WAIS-R 成人知能検査法 日本文化科学社
- 杉村 健・伊谷 実 1991 中学生における生活実態調査, 教心総会, 749-750.
- 鈴木規夫・山田文康・池田輝政・赤木愛和 1991 国語の試験問題の出題形式に関する比較研究 大学入試センター研究紀要, 20, 1-45.
- 相談と検査の研究会 1991 SG 式進路適性検査 実務教育出版
- 孫 媛・芝 祐順 1990 特異な反応パターンを示す被験者の能力推定 ―一般項目反応理論の適用― 教心研, 38, 360-368.
- 鈴木 学 1983 個人差をもう少し考慮した潜在特性モデル 第11回行動計量学会大会発表論文集, 12-14.
- 竹谷 誠 1991 新・テスト理論 早稲田大学出版部
- 東京大学教養部統計学教室(編) 1991 統計学入門 東京大学出版会
- 豊田秀樹 1990 共分散構造の表現 教心研, 38, 438-444.
- 豊田秀樹 1991a 尺度値間の距離が反応確率を規定する項目反応曲線, 心研, 61, 377-383.
- 豊田秀樹 1991b CALIS プロシジャによる共分散構造分析入門 日本 SAS ユーザー会論文集, 79-90.
- 豊田秀樹・山村 滋 1991 数学の試験問題の出題形式と設問過程に関する比較研究 大学入試センター研究紀要, 20, 73-91.
- 豊田秀樹・柳井晴夫・前田忠彦 1991 大学入試改善に関する相反する意見の比較研究 教心総会, 753-754.
- 土田昭司・二宮智子・平野英一 1991 住宅ローンが, 個人の経済判断と消費行動の関係におよぼす効果 ―CALIS による多重指標分析の使用例 日本 SAS ユーザー会論文集, 97-100.
- 辻平治郎 1991 パーソナリティの5因子説の検討とその測定尺度作成の試み, 日心大会, 601.
- 植田 智・吉森 護 1991 日本版 MLAM 承認欲求尺度作成の試み 広島大学教育学部紀要 1部, 39, 151-156.
- 梅永雄二・前川久男 1991 WAIS-R による知的障害児・者の知能構造分析 教心総会, 759-760.
- 山田文康・岩坪秀一 1991a 試験問題の定性的評価に基づく科目難易度の推定 ―社会・理科についての第1次報告― 大学入試センター研究紀要, 20, 205-277.
- 山田文康・岩坪秀一 1991b 試験問題の定性的評価に基づく科目難易度の推定 ―社会・理科についての第1次報告― 日本行動計量学会第19回大会発表論文集, 64-65.
- 山口 勸 1991 集団主義的傾向と年齢, 創造性, および他の性格特性との関連について, 教心総会, 765-766.
- 山本嘉一郎・吉村 英・竹村和久 1991 パソコン SPSS 東洋経済新報社
- 柳井晴夫・前川眞一・豊田秀樹・鈴木規夫・仙崎 武・赤木愛和・中島直忠 1991 高等学校における進路担当教師を対象とした進学指導の実態に関する調査研究 ―学力偏差値を主とした進学指導の改善を中心として― 大学入試センター研究紀要, 20, 93-166.
- 柳井晴夫・前川眞一・豊田秀樹・鈴木規夫・前田忠彦 1991 高等学校における進学指導の実態に関する調査 ―特に斜交プロマックス法に基づく2次因子分析の結果を中心にして― 日本行動計量学会第19回大会発表論文集, 60-63.
- 吉澤 正 1991 多重配列データとその解析 統計(日本統計協会), 6, 35-43.