

教養教育段階におけるテストに関する授業開発と実践（1）

○木村 拓也（九州大学基幹教育院），西郡 大（佐賀大学アドミッションセンター）

1. はじめに

木村(2010)で論じたように，米国の教員養成制度を模して始まった，我が国の戦後教員養成において，当初，教育測定が教職教養に位置づけられていたにも関わらず，*Educational Measurement*が「教育統計学」と測定を外して訳出され，1954(昭和29)年の教職教養の単位減(20→14)において，「教育統計学」は必修単位から大学が独自に開講して選択科目として加えるその他の科目になり下がり，1990(平成2)年の改革以降は，選択科目としての記述すらなくなっていったのが，教員養成を巡る「テスト」の取扱いである(木村2010:42-44)。こうした状況から言えるのは，例えば，テストに近いとされる教育学部であったとしても，教員養成に関係が薄い関係上，専任教員をおくことは稀である。勿論，理系も含めて，テスト関係の研究室が存在するところは少なく，大学で「テスト」を学ぶ状況はかなり困難であるということである。実際に，2008年に行ったテスト学会員調査(木村2009，2010)でも，回答者数の半数ほどが「古典的テスト理論」や「項目反応理論」を講義等ではなく，独学で修得したと回答している(木村2010:35)。

テストを教えるカリキュラム開発という観点では，柴山(1996)，松井・柴山(1997)の研究が挙げられる。そこでは，教員養成課程の制約の中で，「テスト学教育」を如何に行うかという科目構成・配置について論じられている。一方，筆者らのように，1999(平成11)年に，東北大学，筑波大学，九州大学にアドミッションセンターが設けられて以降，「テストの専門家」として，各大学のアドミッションセンターに奉職するケースも出てきた。そこでの業務は，アドミッション業務とともに，教養教育段階の講義を担当することが多い。そこで，まず，筆者の一人である木村が，教養教育段階における「テスト学教育」を開始し，その後，同じ研究室出身の西郡を誘い，講義開発と実践を行ってきた。本研究は，二人の共同開発研究の結果である。

仮に，大学で「テスト」について学ぶ場合でも，専門科目における教育測定論（あるいは心理測定論）や教育評価などで学ぶことが一般的であろう。その受講者は，教育学や教育心理学，心理学，教育工学等に関係する分野の学生たちが中心である。

一方で，テストや評価等に関しては，多くの人々が「受け手」として経験している。また教養教育を受講する学生は直前に入試を体験しているため，テストに対して，肯定的・否定的を問わず，印象が新鮮である。にも関わらず，「作り手」の視点からテストを科学的観点からとらえ直す機会は皆無と言っても過言ではない。

そこで筆者らでは，一部講義内容・スライド・ビデオ等を共有，時には，出張講義をしながら，共同で授業開発を行った。両者に共通するのは，「テスト学教育」をテスト理論のみに焦点を当てず，社会の中での「テスト」の位置付けまでを講義対象としたことである。テストの基本的考え方（信頼性や妥当性，テスト理論等）から，テストの歴史/社会学/心理学など様々な観点からテストについて講義することでテストが学問的に検討に値する分野であることを知ってもらうため，「テスト」をテーマとした教養教育科目を開講してきた。

本発表では，木村が担当した京大，長崎大，九大での講義と，西郡が担当した佐賀大での講義を対象に，受講者に対して行った質問紙調査の結果をもとに彼らの，(1)では，「テスト観の変化」・(2)では，「意識の変化」について検証した結果を報告する。表1は，筆者らが，これまで各勤務大学で開講してきた「テスト学教育」関連講義の受講者数であり，これまで約800人近い学生に教授してきた。専門教育よりも教養教育段階の方が，受講者数が多いことが予想され，「テスト学」の広報的意味合いも含意することが可能であった。

表1. テスト関係学講義受講者数

	京大	長大	佐大	九大	年度計
2008	90				90
2009	87				87
2010	9*	24	77		110
2011		83	43		126
2012		102	86		188
2013			53	29	82
2014				92	92
2015				24	24
合計	186	209	259	145	799

* 但し，この年度のみ教育学部専門科目「比較教育学講義-学力測定政策の日米比較」として開講した。

2. 授業の内容

本授業は，2008年度から2015年度にかけて実施してきたものである（2008・2009年度は「テストの科学とその歴史」，2010～2012年度は「社会と歴史—テストの科学とその歴史」，2013～2015年度は「テスト学への招待」として開講）。シラバスには，「入学試験・就職試験・資格試験・昇格(昇任)試験等々，人生において幾度も直面する「テスト」でありながら，「テスト」に関する科学的な知識に接する機会は殆どない。そこで，本講義では，皆さんがいままで当たり前のように受けてきた「テスト」を哲学・歴史学・法律学・社会学・心理学・数学・統計学といった大学諸学問の観点

から分析し、更に、テスト理論と呼ばれるテスト評価測定技術についての導入的な解説を行う。「テスト」を単に「害悪」と捉えるのではなく、「テストの結果が、個人の処遇や人生を大きく左右するものであるか故に、その実施にあたっては、細心の注意を払うべき類のものである」との認識に立って、よりよい「テスト」を実施していくための「基礎教養」の修得を目指す」という概要を記している。

授業の内容は、表1の通りである。前半を人文社会科学から見たテスト、後半を統計科学から見たテストとテーマを区切り、それぞれ、テスト学文献リストを配布し、その中から1冊選んで書く読書レポート、仮想テストデータを配布し、古典的テスト理論で分析をするデータ演習レポートを課している。

講義内容は、「テスト学の基礎」では、身近なテストを思い出してもらいながら、CBTなどの現代的なトピックを交え、テスト作題手順などについて基礎知識を教授する。「テストの社会学」では、メリトクラシー論を通してみたテストのあり方を講義し、「テストの哲学」では、アフーマティブ・アクションなどの米国のテストを巡る裁判事例、人種分離統合教育の歴史を開説した上で、マイケル・サンデル教授の正義論の講義を見ながら、社会哲学における分配の公正原理からみた大学入試の社会的役割を考察する。また、同時に、NHKスペシャルの「SATの真実」というドキュメンタリーも同時に活用している。次に、「テストの歴史学」では、江戸時代の試験制度を概観し、現代のテストの異同を確認し、「テストの心理学」では、公正感からみた面接テストの設計について講義する。更に、「テストの法律学」では、調査書や指導要録を巡る裁判事例や内申書裁判の事例からテストの法律学が問題にしてきた論点について解説をして、前半を終える。各回とも、テストという共通テーマに各学問の立場から見ることで、各ディシプリンの視角を意識的に話すよう心がけている。

表2. 講義内容 (2015年度シラバスより抜粋)

第1回: オリエンテーション
第2回: テスト学の基礎—テスト学の最前線, 暗黙のルール
第3回: テストの社会学—学歴社会の理論, メリトクラシー論
第4回: テストの哲学 (1) —アフーマティブ・アクション
第5回: テストの哲学 (2) —正義論から見たテスト
第6回: テストの心理学—社会心理学から見た入試の公平感
第7回: テストの歴史学—江戸時代の試験とSAT
第8回: テストの法律学—教育法体系における試験, 裁判事例
第9回: テストの数学—合計得点・二段階選抜・合否決定
第10回: 統計学の基礎 (1) —偏差値
第11回: 統計学の基礎 (2) —相関係数
第12回: テストの統計学 (1) —項目分析と統計的方法
第13回: テストの統計学 (2) —信頼性・妥当性
第14回: テスト現場の実際 (1) —人事アセスメントの考え方
第15回: テスト現場の実際 (2) —M-1グランプリの信頼性

* 下線は、著者間で共有している講義コンテンツ

また、後半では、合否入替り率の図などを参考に、合計得点の二次元平面での仕組みを解説することに始まり、標準偏差と分散についての解説から偏差値の紹

介、相関係数の解説を行い、テスト理論の理解に必要な不可欠な基本統計量の理解を深めた上で、データ演習レポートで行う、項目分析の各種道具立てや信頼性係数の算出方法、折半法などの考え方を解説する。

その後、テストデータの実際の分析方法、信頼性・妥当性の考え方を具体的に示せる意味で、人事アセスメントの考え方とM-1グランプリの得点データを用いた独自の分析結果を例示として解説し、古典的テスト理論の道具立ての実際の活用場面の理解を深めている。

3. 質問紙調査の結果

筆者らは、講義内容が固まり始めた、2012年度から互いに相談をしながら、当該授業の受講学生を対象とした質問紙調査を行っている。本発表で分析するデータは2012年度の102名(長崎大)、2013年度、2014年度の121名(九州大)の計223名であり、いずれも同じ講義資料を使っている。

3.1 単純集計

「テストの印象」について、4件法で「非常に良い」「まあまあよい」「あまりよくない」「非常に悪い」を事前事後に問うた結果、189名中、印象が向上したのは106名(56.1%)。印象が低下したのは13名(6.9%)、印象が変化しなかったのは70名(37.0%)であった。

次に、「テストの必要性」について、2件法で「必要／不要」を尋ねたところ、189名中184名(97.4%)が初回も講義終了時も必要と回答している。初回で「不要」講義終了時で「必要」回答したのは4名(2.1%)、初回で「必要」講義終了時で「不要」と回答したのは1名(0.5%)であった。

更に、「テストの公平性」について、2件法で「公平／不公平」を尋ねたところ、188名中66名(35.1%)が初回「不公平」で講義終了時も「不公平」と変わらず、また、初回で「公平」で講義終了時に「不公平」と回答したのは、78名(41.5%)である。逆に、初回で「不公平」で講義終了時に「公平」と回答したのは、14名(7.5%)、初回から講義終了時まで「賛成」であったのは30名(16.0%)であった。

最後に、講義を通して、「テスト観が変わった」と回答したのは、187名中174名(93.0%)であった。

3.2 テキストマイニングによる印象変化の測定

以下、自由記述部分を中心に、テストイメージ等に関するテスト学教育の効果測定について、多重対応分析した結果を報告する。

3.2.1 「テスト」イメージの事前事後の関係性

テスト学教育を行う前後に、「『テスト』と聞いて、思い浮かぶイメージを思いつく限り全て挙げて下さい」と尋ねた結果が、図1である。単純に見て、テスト学教育を行った後の方が当たり前であるが、テスト理論固有の専門用語が増えている。第一象限には、その専門用語が羅列されてあるが、特に、事前のイメージとの関係性はないようである。第二・第三象限では、事後に、「妥当性」「信頼性」「公平性」が挙げられており、

事前のイメージとして「面接」や「受験」が上がっている。事前の入試を通したテストに関する興味関心から、テスト理論の考え方を身につけた学生群であると思われる。第二象限の左端で、事前に「心理テスト」のイメージをもっていた学生が、事後のイメージで「妥協」「完璧なテストは存在しない」「なかなか難しい」などの印象を持っていることも興味深い。

3.2.2 「テスト学」講義の印象と「テスト観」の変化理由の関係性

次に、「テスト学」講義で最も印象に残ったことと「テスト観」の変化理由の関係性について見た結果が、図2である。第一象限では、講義の印象が「様々な視点」「見方が変わる」「項目分析」「信頼性係数」他、基本統計となっており、変化理由が「テスト科学」「奥深い」「テスト観」「分析が必要」など、テストデータの分析を通して見えた新しい世界観に惹かれた様子が伺える。第二象限では、講義の印象、変化理由ともに「信頼性」「妥当性」「 α 計数」であり、テスト理論の重要性を認識した学生群であることがわかる。この層がテストの最終的な印象が「非常に良い」学生群であることも特筆すべきであろう。第三象限では、講義印象がテストの「公平性」「採点基準」「M-1」「人種」であり、変化理由も「テストの公平」であることから、テスト学が追求する「公平性」の観点に関心を持つ層であると同える。

3.2.3 「テスト学の真髄」の規定要因の探索

最後に、「テスト観」の変化理由と、講義で感じた「テスト学の真髄」についての内容把握について見た結果が図3である。ここでは、質問項目の関係から九州大学のみのデータ(N=121名)である。右側の第一象限・

第四象限には、真髄の把握として「公平性の追求」「誤差を考える」「信頼性/妥当性の向上」「ベストではなくベターを目指す」等が有り、変化理由として「テストの科学」「奥深い」「信頼性/妥当性」等があり、テスト理論に関する内容が併置されている。第二象限には、真髄の把握として「完全なテストは存在しない」であり、変化理由は「知って驚いた」「テスト観」「新たな観点」「テストの公平」が挙げられており、完全なテストが存在するという前提が覆された学生層であることが伺える。第三象限には、真髄の把握で「テストの分析」「学問を感じる」であり、変化理由の「分析が必要」とのことからデータ分析に興味を惹かれた層、また、真髄の把握が「社会への寄与」と変化理由の「作る側」が併置されており、作成面側からみた時に、社会貢献を感じた層が存在することが伺える。

3.3 「テスト学教育」の得点分析

従属変数を講義成績(総得点、読書レポート成績、データ演習レポート成績)にし、独立変数をダミー変数として、文理の別や、テキストマイニングで得られたカテゴリー(「変化理由」「真髄把握」)を投入して重回帰分析を行った。その結果、読書レポート(自由度調整済 $R^2=.256$)では、「真髄_学問を感じる」($\beta=.363^{***}$)、「理系」($\beta=-.245^{***}$)、データ演習レポート(自由度調整済 $R^2=.164$)では、「初回_公平性」($\beta=-.242^{**}$)、「真髄_ベターを目指す」($\beta=.209^{**}$)、最終成績(自由度調整済 $R^2=.171$)では、「真髄_学問を感じる」($\beta=.312^{***}$)、「真髄_ベターを目指す」($\beta=.248^{**}$)となっている。学問としてテスト学を認識した学生の成績が高い傾向にあることが分かる。

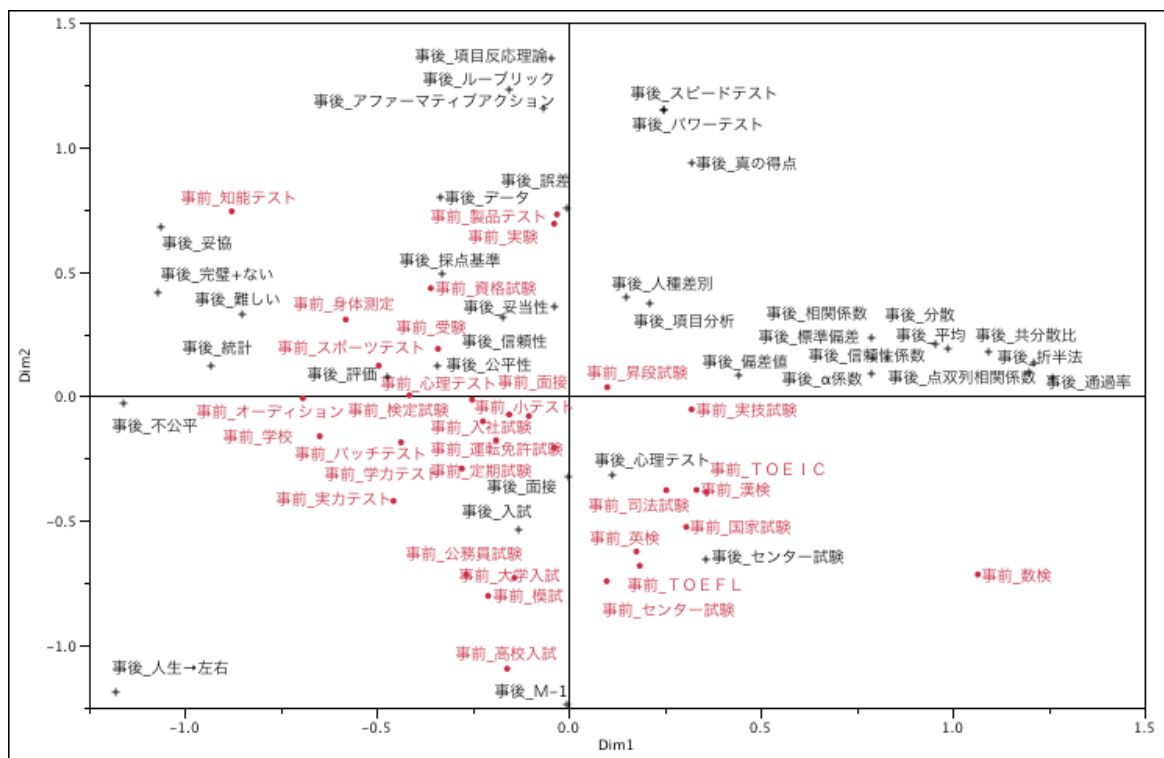


図1. 「テスト学教育」受講前後におけるテストイメージの変化

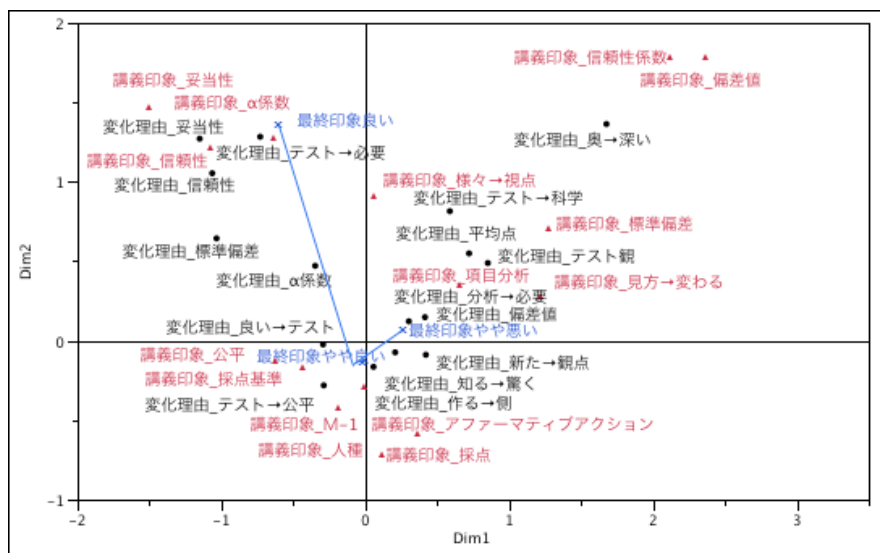


図2. 「テスト学教育」受講前後における「テスト観の変化理由」と「講義で印象に残ったこと」の関係性

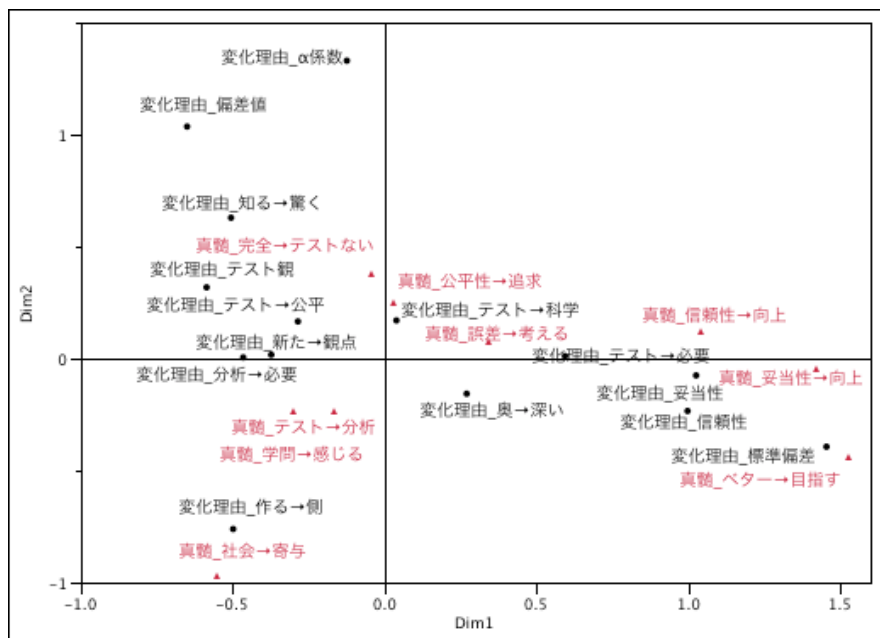


図3. 「テスト学教育」受講後における「テスト観」の変化理由と「テスト学」の真髄把握の関係

4. まとめ

筆者の一人(木村)が、まだ、教育哲学・思想史の院生であった頃、伝統的な教育学においては「テスト・試験」は批判すべき対象であれ、科学的知見から研究する対象であると想定していなかった。たまたま他専攻の専門科目である「教育測定」(野口裕之先生)の講義を受けた時の、こんな学問があるのだ、と感じた衝撃は今でも鮮明に覚えている。そんな経験があるが故に、テスト学講義が存在することの重要性を感じている。

その後、筆者らは、テストの世界に足を踏み入れることになるが、教養教育段階でのこうした授業の中から、もし将来、テスト学研究を志してくれる学生が現れたとしたら、筆者らにとって望外の喜びである。

2015年度より、九州大学大学院人間環境学府教育システム専攻において、教育テスト論研究室を立ち上げ、

学生指導が可能となった。教養教育段階での授業経験を活かし、次の段階では、テスト理論のみならず、テストの社会的背景までもを含めた、学生指導・論文指導を行っていきたいと考えている。

【文献】

- 柴山直 (1996) . 「テスト理論のための授業カリキュラム」『学習評価研究』 27, pp.44-51.
- 松井仁・柴山直 (1997) 「教員養成課程における統計・測定教育試論」『大学教育研究年報』 3, p.24-31.
- 木村拓也 (2009) . 「『テストの専門家』を巡る人材養成の現状と課題——テスト関係学会調査から見えてくるもの」『人事試験研究』 212, pp.2-15.
- 木村拓也 (2010) . 「日本における『テストの専門家』を巡る人材養成状況の量的把握」、日本テスト学会編『日本テスト学会誌』 6, pp.29-49.