

第4回講演会

「新たな時代における資格・検定試験のあり方」 レポート



平成 25 年 10 月 4 日(金) 午後 1 時 00 分～午後 4 時 30 分
東海大学校友会館 望星の間

信頼と安心の検定を目指す

全 検
ZENKENKIKOU
機 構



去る10月4日、今回から文部科学省後援となった全国検定振興機構による講演会が、霞が関ビルにおいて開催されました。

この講演会は、検定試験の振興と、検定業界における情報交換を目的に開催され、今回で4回目。毎年、資格検定や認定に関わる事業者、研究者、メディアの方々をお招きし、さまざまな視点での検定をめぐるお話を伺っています。

今年度のテーマは、「新たな時代における資格・検定試験のあり方」。会場には、事前に申し込みをされた100人近い方がつめかけ、熱心に聞き入ったり、メモを取る姿が印象的でした。

講演は、当機構理事長・池田央の挨拶からはじまり、続いて、文部科学省生涯学習政策局 生涯学習推進課課長・早川俊章氏から、文部科学省のこれまでの資格・検定試験の取り組み、機構への今後の期待などのお話がありました。

基調講演では、昨年の講演に引き続き、テスト理論などで教鞭を取り、項目応答理論などで数々の論文のある東北大学教育情報学研究部・教育部教授 村木英治氏が、テスト作成におけるバイブルともいえる『教育・心理検査法のスタンダード』に準じ、テストの作成の方法と重要点を解説。その裏付けともなる複雑な数式には、聴講者からはうなり声も聞こえましたが、本レポートではその抄録版をお届けします。

休憩をはさんでからは一転、NPO教育支援協会で理事を務める吉田博彦氏と、元・文部科学省で、現在独立行政法人 国立高等専門学校・理事の上月正博氏による、民×官の対談バトルが繰り広げられました。両名における、今後の民間と官の協力体制はかくあるべきという忌憚のない討論に、会場からは笑い声もしばしば。

また、会場では休憩時間などに、講演者と、あるいは、聴講者同士がおたがいに名刺交換をするシーンも見られました。情報交換と交流という本講演会の目的が根付きつつある。そんなことを考えさせられた1日でした。



文部科学省生涯学習政策局
生涯学習推進課課長
早川 俊章 氏



東北大学教育情報学研究部
教育部教授
村木 英治 氏



NPO教育支援協会
代表理事
吉田 博彦 氏



独立行政法人 国立高等専門学校機構
理事
上月 正博 氏

来賓挨拶



「生涯学習と検定試験への取り組みと、 よりよい検定試験へ向けて」

文部科学省生涯学習政策局
生涯学習推進課課長 早川俊章氏

検定試験の
評価ガイドラインと
自己評価シートの普及
およびその調査結果について



文部科学省生涯学習推進課の早川と申します。全国検定振興機構の第4回講演会にあたりまして、一言ご挨拶を申し上げます。

全国検定振興機構におきましては、平成21年の設立以来、一貫して検定業界全体を見据えて、検定試験の実施団体に対します支援活動にお取組みいただいていることに、心より敬意を申し上げる次第です。

言うまでもありませんが、検定試験は、チャレンジ精神の涵養(かんよう)や、自己の到達度の確認、また、継続的な学習意欲の喚起や向上などなど、実にさまざまな意義を持っています。これらはまさに、個々人の学習成果を適切に評価し、そしてそれを生かすことのできる社会、いわゆる生涯学習社会の実現においても、大変重要なものであると考えています。

こうしたことから、文部科学省では、平成22年に取りまとめた『検定試験の評価ガイドライン』や、平成23年に民間指導で作成していただきました『自己評価シート』の普及・啓発を行いながら、検定事業の自己評価、あるいは情報公開の推進に取り組んでまいりました。

また、今年の6月ですが、第2期の教育振興基本計画が閣議決定されましたが、そのなかでも検定試験の社会的な通用性を高める観点から、検定業者における自己評価あるいは情報公開の取り組みをさらに普及させていくということが、目標として掲げられました。

同じく6月ですが、文部科学省でも、検定事業者における情報公開、あるいは自己評価の実施状況について、調査を行い、結果を公表しておりますので、今後のお取組みに参考ご活用いただければと思っています。

本日は、「新たな時代における資格・検定試験のあり方」をテーマに、基調講演、対論が予定されているとかがっています。今後の検定試験のあり方を考えるうえで大変意義深いものであると考えておりますので、今後も、より魅力的な検定試験の実現に向けてご尽力いただければ幸いです。

最後になりますが、この講演会の開催にあたりまして、種々ご尽力いただきました全国検定振興機構・池田理事長さんをはじめ、関係の皆様にご感謝を申し上げますとともに、本講演会の成功と、検定業界全体の、今後ますますのご発展を祈念いたしまして、ご挨拶とさせていただきます。

どうもありがとうございました。



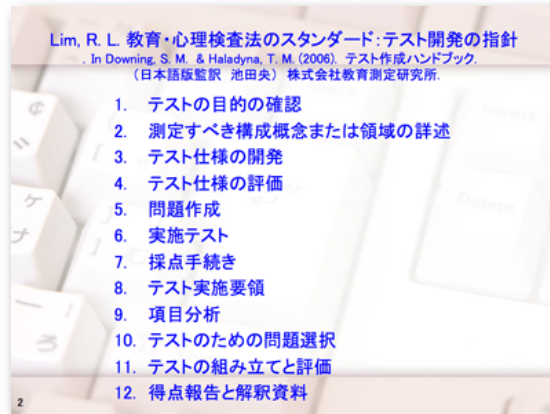
「Psychometrics (心理統計学) が教える テスト作成プロセス：Paper 版と Computer 版」

東北大学教育情報学研究部
教育部教授 村木 英治 氏

なぜそのテストを行うのか
妥当性を問い続けることが
テスト作成の
第一歩である

東北大の村木です。去年は、CAT・コンピューター関係のテストをお話ししたのですが、今回は、心理測定学の基礎的な部分をお話ししたいと思います。タイトルが心理統計学となっていますが、Psychometricsには測定、統計など、いろいろなものが含まれるんですね。

さて、今回は『テスト作成ハンドブック』（教育測定研究所）の、第2章を元にいろいろ話していきたいと思っています。この本は、割とコンパクトな形で、テスト作成のあれこれを網羅しています（図1）。



ここでは(テスト問題の作成、実施について)、このとおりやっていけばいいというのが示されています。ただ、これをリニアにやればいいというのではなく、繰り返しでやっていくべきものと思いますね。まずは、最初の、テスト目的の確認というところからお話させていただきます。

テスト目的の確認というのは、なぜそのテストを行うのか、基礎中の基礎になります。テスト開発の論理的な出発点であり、スペック作成のための大事なものの、基本精神ですね。その中には、テスト内容の決定、形式、応答のしかたとか、妥当性の問題、信頼性の問題、あるいはフェアネスというのが関係してきます。

さて、資料の中にも、妥当性と、信頼性という言葉がよく出てきます。妥当性については、「効果的なテストを作るには、テスト得点からなされる推論を支持するのに十分な妥当性が確保できるように体系的で、うまく組み立てられて



方法が必要である(Downing, S. M.)」などのように、たぶんみなさん聞かれると思うのですが、(具体的には)何を言ってるのか分からない。つまり、これらを操作的な定義に置き換えなければならない。そのなかで、さまざまな妥当性に関する検討があったわけですが、1985年のスタンダードでは妥当性を「テスト得点を用いたある得点の推論が、適切であるか、意味が有るか、有用であるかを示す」と定義し、「テストの妥当性とは、かかる推論を裏付けるための証拠を累積するプロセスであるとしています。他方、妥当性は心理学的概念ですから、仮定を確かめる調査、研究=テストをめぐる研究の方法のすべては、妥当性検証のプロセスであるというコンセンサスが定まってきました。つまり、テストの妥当性とは、(そのテストが)「こんなことができるんじゃないか、あんなことに役に立つんじゃないか」といったいろんな推論を確かめ、それを証拠固めしていく普段の努力、研究、検証であるということが、僕たちスペシャリストの間で、だんだん成立してきました。

一方で、信頼性というのは、数値計算で出てきますので楽なんです。テストの信頼性とは、テスト結果の正確性あるいは一貫性(安定性)を表す概念なんです。そこには、測定誤差も含まれてしまいます。それは、試験を受ける側の気分や体調といった各個人内のランダムな変動や、状況要因などが含まれます。古典的テスト理論では $X=T+E$ (真値)+E(誤差) で表しますが、これを発展させて信頼性係数を表す公式が存在します。ものの本によれば、信頼係数は0.8以上が望ましいとされていますが、たとえば、信頼性係数が0.5のテストを0.8に引き上げるためには、あと何題問題を追加すればいいといったことが、この式から求められます。

次に、信頼性係数と妥当性の相関ですが、基本式から、次のことが結論付けられます。それは、つまり、妥当性はものすごく大切なもので、妥当性を検証しなければ、テストは成り立たないわけですが、その妥当性には、上限があるということです。妥当性係数は、信頼性係数の平方根を上限とするということです。妥当性を高めようとしても信頼性係数が低かったら、妥当性は高まらないということです。考えてみればあたりまえの話で、信頼性が置けない問題のテストをいくら作っても、何の役にも立たない、ですね。ただ、信頼性が高ければ妥当性が高くなるかといえばそうではない。不明です。しかし、信頼性が低くなれば、妥当性は自動的に低くなってしまいます。テストの目的に合致した妥当性は大切ですが、それに付随して、信頼性も大事だということです。

次に、測定すべき構成概念の話です。ひとつは、無関係な分散、つまり、測ろうとしていないものを拾ってくることをなくそうという見直し、そして、測ろうとしているのに、それをテスト項目の中に反映していないという見直し。これを代表性不足といいます。それらについて書かれています。テストの目的、テストが測ろうとする構成概念、あるいは内的な属性といったものを、テストの枠組みに拡張しようという仕事、それが、テストのスペシフィケーションです。ですから、テストのスペックを作る場合は、なるだけ詳しく書いていくということが必要になっていくわけですね。

さらに、テスト仕様の開発ということで書かれています。問題、質問の形式、あるいは解答にあたっての条件や、採点手続きといった詳細を、テスト仕様の中に盛り込んで書いていき、どれだけテストの内容をカバーしているか、範囲をはみ出してないか、あるいはそのテストで測る能力をイヴォークしているかが大事だと書かれています。

最後に、項目のプールデザインについてお話しします。新しくCAT(Computer Adaptive Test)になる場合に、このテスト項目を使った場合、次のテスト項

CATで複雑化する
プールデザインに合わせ
求められる
サイコメトリックスの進化



目は使えないといった場合もありますね。逆に、これを使ったら、次にこれを入れたいといった場合も出てきます。これが、CATにおける項目間の相対的な位置関係と呼ばれるものです。資料ではペーパーテストの内的属性をリストしましたけれど、コンピューターのテストになりますと、新しいサイコメトリックスな様々な考慮しなければいけない状況が生まれてくるというのが確かなようです。サイコメトリックスもテストの発展とともに、進んでいかなければならないということです。

質疑応答から

Q

専門外なので、正直なところ難解なお話でしたが、お話の内容を理解するためのテキストなどを提示いただけませんか

A

アメリカのほうで盛んな学問ですので、アメリカのテキストブックというのがスタンダードになります。そういう意味では『テスト作成ハンドブック』は池田理事長が中心となって私どもが訳しましたので読んでいただければと思います。また、私の本の中にも網羅していますので、そうした文献についても当たっていただければと思います。





対 論

「新しい公共の時代における民間教育機関の使命」

NPO 教育支援協会代表理事

吉田博彦氏

×

独立行政法人国立高等専門学校機構理事

上月正博氏

なぜ

『検定試験の評価ガイドライン』

は生まれたのか。

検定における質は、

だれが保証するのか



吉田 「吉田です。今回お話を受けたときに、村木先生の後に講演するのでは参加された方々が死んじゃうよということで、講演はやめようということになって(笑)。『数学の時間』の後は『総合的学習の時間』ぐらいがちょうどいいんじゃないかということで、それでは全検が出来たときの推進課の課長をやられていた上月さんにお話を聞こうということになりました」

上月 「今日は、6年前に文科省の生涯学習推進課ということになっておまして、『検定ガイドライン』というのを作ったのを思い出しております」

吉田 「早速ですが、今日のテーマを、今流行の半沢直樹に例えると、『民間でしっかりやっているのに、これ以上どうしろと言うのか!』という感じだと思うのですが、官民の対立構造になっては仕方ないので、協力体制を組むにはどうすればいいのかといったことを前提にお話をさせていただきたいと思います」

上月 「文科省当時、学んだ成果を適切に生かすことができる社会の実現のために、なんらかの手立てを打っていくのが、ひとつの大きな仕事かなあと、そういう認識で捉えていました。それ以前ですが、学校教育で、これは、マスコミなどを含めてですが、何かが起きると『教えるべき』という“べき論”が持ち上がり、何か違和感を感じていた部分もあったのですが、どうやったら、本当に必要なことが習得され本人が意識をされ、社会的にも意識をされるのかといったことが、個人的にも関心事だったんです。そんな中で、検定事業の仕事があったわけですが、そこでびっくりしたのは、検定によっては何百万人も受けるものがあったんです。これはもしからしたら、教える側主体のシステムから学習者主体のシステムに、少し支援できないかと思ったんです」

吉田 「そのころ教育基本法が改訂され、生涯学習の機運が盛り込まれたわけですが、当時課長をされていた上月さんからすると、民間で行われている検定事業が、相当社会的な影響力を持っていたにも関わらず、かなりいいかげんにやっているとところもあるんじゃないか、そういう気持ちは文科省の中になかったですか？」



上月 「まさに玉石混交で、民間がやるということは需要も高まっていいという議論に合わせ、自由度を尊重しつつ質をどう確保していくのかは、大きなポイントだなと思いましたね。2000年代の前半には、小泉改革で、できることは民間にという機運があったのですが、後半に入ると、『それでいいのか』という雰囲気もあったんです。検定団体の不祥事や問題もありましたし、継続性が重要となる学習分野でなかなか簡単な問題じゃないなという感覚が出始めたころです」

吉田 「そんななかで、民間でやっている全検構みたいのができてきたわけですが、上月さん自身は、全検みたいのができるのが望ましいと思っていました？」

上月 「正直、いろいろ考えました。大学は、大学評価を受けるわけですけど、その機関を文科省が認定する。しかし、それには大規模な組織を作らないといけないし、そこまでの余力はないので、大手の団体に相互評価をする話などもしましたね」

新しい公共の時代
大切なのは、
官民で目的を共有化すること



吉田 「もし、学習評価を国が担うとすると、プラスとマイナスの部分があると思うんですが、上月さんはどう考えます」

上月 「生涯学習は学校から全部入るわけなんで、評価に対する信頼度をどこまで確保するかということとのバランスだと思います。例えば、日本の大学入学資格は高校の校長先生が卒業認定すれば終わりなわけですけど、国家が関与してやっているところも少なくない。また、職業訓練の世界ではアジア各国で、質の確保のための認証システムをお互いに作りましょうということをやっている、それが主流になりつつあるので、行政がなんらかのフレームワークに関与して間接的には保証するのがいいのかなとは思っています。それと、テストをどこで実施するのかという問題もあります。

きっちりとしたテストには、ものすごいコストがかかるんです。

推進課長のときに、中学卒業認定試験を所管していたのですが、これは国が直接実施しています。高卒認定試験もそうなんですが、正直、けっこう大変なので、民間でやってもらえればと思ったこともあります。

ただ、これを民間にするとコストが受験料に反映してくる。そこがバランスなんだと思います。また、国が関与しているということではセンター試験があります。本来、大学が独自に入試テストをすべきだとも思うのですが、それにはコストがかかり、しっかりした試験をやろうとすれば、さらにコストがかかるため、現在の大学にはそんな余力はないといった問題もあります。教員の方も、余計な仕事と思ってる方も多いんですね」

吉田 「大学で入試問題を作っている先生には高校の学習指導要領の中身を知らない方もたくさんいるじゃないですか。それでは問題を出される生徒のほうが大変ですよ……。さて、次に、テストの信頼性ということですが、私は、信頼性というのは客観的な数値で表せると思っています。多くの方は国がやることで信頼性が高まると思ってますが、私はむしろ行政などがやるほうが信頼性がないと思うんだけど、そのへんはどうなんですかね」

上月 「確かにそういうことはあるかと思っています。2008年にですね、『検定試験の評価ガイドライン』というのを出しました。趣旨は質保証なんですけど、それを参考にしてほしい。もうひとつは、それを元にして、民間で



第三者評価組織を作って評価する。そういうのがいいかなと思って作ったんです。当初は、行政が関与するわけでもないのにそうしたガイドラインを出すことが関係者から理解されなかったのですが、大手をはじめ採用する事業者さんもあり、評価組織に関する話も聞きますので、よかったのかなと思っています。それと、当時思ったのは、質保証のガイドラインがあるわけですから、もし、検定への後援名義を求められれば、ガイドラインを基に考えますよ、ということです」

吉田 「それは基準がしっかりしていれば、民間がやってる検定とか認定に対して文科省の後援がありうるということですか」

上月 「そうです。ただしコール条件じゃないですよ。それをひとつの大きな要素にするということです。ただ、それは段階的な話で、さきほどのような第三者組織が認定していくのが、将来的には一番いい姿だと思います」

吉田 「行政がやることをちゃんとやってきて、民間は、今回の全検ではないですが評価基準を自分たちで作るなどやってきて、で、それを担ってくれる団体がいくつかあって、それでタッグを組めれば、うまく動くと思うんですよ。半沢直樹じゃないですけど、民間には、行政っていうのは検査だの査察に入っているというイメージが強いんですよ。だから、“一緒にやる”ってことを発信できれば、大分違ってくるとは思うんですよ」

上月 「目的を共有化するということですよ。役割はもちろん違うんだけど、役割をお互い尊重して目的を共有できれば、非常にうまくいくことが多くなりますよね」

吉田 「僕はよく役所の方に言うんだけど、自分たちで飯を食っているのだから民間はお行儀悪いですよ。それを『お行儀悪い』って言われちゃうと、前に進まなくなっちゃう。なので、これから改善が必要な部分もあることも前提にして、文科省にも理解してもらって組んでいければいいと思うんです」

上月 「生涯学習政策局っていうのは、予算も少ないですし、権限もほとんどないんです。そうすると、民間の方と一緒にやらないと、政策のミッションを達成できない立場なんです。さっきも言ったように、目的が共有できれば、民間の方もいろいろな方法論を持っていますので組んでやっていけると思います」

質疑応答から

Q

現在、厚生労働省が幅広く認定、指導というのを
行っています。
そういうものとの関係をどう考えますか。

A

あまり詳しくはないですが、厚労省は、安全性や職業能力の仕組みを持っているので、その関係で出しているのだと思います。一時は我々もそうした、個別認定をする方法も考えましたが、労力に対し、効果が少ないと判断した経緯があります。そうしたこともあり、ガイドラインというものが生まれたのですが、それに沿ってもらえれば、皆様方の事業の信頼度が増すという参考例になったと思っています。少なくとも教育関連部局としては、さきほどの民間との関係で、バランスが取れるのではないかと思います（上