

学生による授業評価に関する物言い

マークシート方式は統計処理には向いているが、
意味ある結果を得るためにはどうすれば良いのか？

もっと工夫を !!

そして、学生自身の手になる評価方式への移行を !!

井上淳 2002-10-08

具体的問題点：後期第1回目の講義で、まず今学期の授業内容を説明した後に、

これで今学期のシラバスを説明したことになる。前学期に実施した「学生による授業評価」をみると、シラバスは役に立たなかったという人が圧倒的に多かった。

ところで、君たちはシラバスを見たのかな？ 挙手してもらおうと、どうも見てもない人が多いようだ。とすると、あのアンケートに、脈絡を考えずに、忠実に答えると、見ていないのなら「役に立たなかった」となるのはもっともである。シラバスを見たか、見なかったか、という項目はなかったのだし、アンケートを取る側が勝手に「シラバスを見たはずだ」という仮定をしているとすると、偽りの仮定から導かれた結論は何を言っても良いというのが論理学の結論では？。

統計処理をしたときにも、意味があるアンケートの取り方、はどうしたら良いのか？ 即ち、(1)でこう答えたら(2)や(3)でこう答えないと意味が判然としないか、或いは全く無責任な答え方をしていると判定でき、それを除外して統計処理する方法はないのか？

シラバスを見たのか見なかったのかはともかく、講義内容や講義のやり方に対する諸君の評価についても私には腑に落ちない点がある。実はこのようなアンケートは91年頃から私や物理の一部教官が始めたものである。私の場合その結果に対して統計は取らなかったが、むしろ自由記述の感想の部分を見てから、アンケート項目に対するその講義に受講している諸君の答えを検討し、その意味を考えたものである。ところが、自由記述の部分はもっとも統計処理に馴染まない。

ところで、2002年度前期微積分第1の中間試験、期末試験に書かれていた感想等は整理してあるので、ホームページに載せる予定である。

理学部地球惑星学科での「学生達による授業評価」をもっと大学として評価し検討したら如何か？

評価の透明性、公平性、客観性をどうしたら保証できるのか？ のもっと徹底した議論をし、「東工大評価方式」とでもいうモデルを少なくとも日本社会に提示する意気込みを！ 文部科学省とか世間に対する御機嫌伺いとして「学生による授業評価をやっている」などという「振り」はしないにしよう！

マークシート方式についての参考資料：TOEICのQ&Aより

Q1：TOEICはマークシート方式を採っていますが、たとえば、選択肢のAだけに印をつければ、4分の1の確率で正解をねらえるのではないのでしょうか。

A1：ご指摘のように、AなりBなりにでたらめに印をつけても、正解となる確率は4分の1で、200問中の50問位は正答になります。しかし、TOEICスコアは正答1問につき5点といった単純な方法で出しているわけではなく、特別な換算表を使ってTOEICスコアを算出しています。たとえ正答数が4分の1であってもそこから算出されるTOEICスコアは4分の1の250点どころか6分の1にも満たないような非常に低いものです。これはTOEICを開発したETSが、それ以前に外国人留学生を対象としたTOEFLというテストを開発し、実施してきた経験等から生み出した独自のノウハウです。一例を上げますと、

TOEICの作成には言語学の専門家に加え、心理学や統計学の専門家が加わりそれぞれの立場からテストを実施した際に起こるであろうと予測される問題¹の解決にあたっており、それだけテストとして完成度が高い

¹ 日本の組織で何かを予測する時、都合が悪い事は起こらないし、それを指摘してはならないと言う暗黙の了解があるのでは？

ものになっている² ということです。TOEIC は、でたために解答すれば、それなりの低いスコアしかでてこないような設計になっているのです。

Q3: TOEIC には Listening と Reading の テストしかありませんが、どうして Speaking テストがないのでしょうか。これ で本当に英語能力が正しく評価できるのでしょうか。

A3: 先ず申し上げておくべきこととして、TOEIC は

その開発に際し、一度に大量の受験者の英語能力を、実施に際して専門家の手を必要とせずに、客観的に、極力低い費用で測定できるような テストとなるようにという条件

が課せられました³。そこで TOEIC の開発にあつた ETS では、TOEIC 実施にあたって予め Listening と Speaking、Reading と Writing との相関関係について検証し、それぞれが非常に高い相関関係を示す ことから、Listening と Reading のみの試験から Speaking と Writing 能力を含めた英語能力が測定できることを統計的に証明しています。そのため、TOEIC は Listening と Reading のみで構成されています。ただしこれはあくまでも統計上の話であって、一人一人の Speaking および Writing 能力を厳密に測るためには Speaking や Writing のテストを行う必要があります。ところが Speaking や Writing は機械採点の可能な客観的ペーパーテストで行える Listening、Reading とは違い人が採点する方式のテストを行う形になります。その場合には、しっかりとした理論的根拠と、それに基づいた専門家による採点者のトレーニングが必要となります。そして当然それなりの費用と確実な理論、それにこれを実行するための管理体制が必要となります。費用や、てまひまを問題とせず、あくまでも厳密な Speaking 能力の測定を必要とする場合には TOEIC と厳密なインタビューテストの併用をお勧めしますが、そうでない場合には TOEIC だけで十分だと思えます。

Q10: TOEIC を受験するにあたって、再受験や事前準備等で著しく有利になることはあるのでしょうか。よく英会話 学校や出版物等で、TOEIC 対策と銘打ったものが目立つのですがいかがでしょうか。

A10: 一般的に、実力以外で高得点を とる可能性があると思われる要因には、多肢選択において推測による得点効果 (Guessing effect)、練習による得点効果 (Practice effect)、受験指導による得点効果 (Coaching effect) の 3 種類があります。このうち「推測による得点効果」については、Q1 の「TOEIC はマークシート方式を採っていますが、たとえば、選択肢の A だけに印をつければ、4 分の 1 の確率で正解をねらえるのではないのでしょうか。」の回答で説明しているように、対策はとられています。「練習による得点効果」は主として、問題形式に慣れることにより得点上有利になることを意味し、「傾向と対策」といった対応があげられます。しかし、TOEIC の形式は毎回同じであり、受験案内等でもサンプル問題といった形で明記されています。開発した ETS は、すべての受験者に事前によく 出題の形式に慣れておいて欲しいといっています。これは、形式に慣れることでスコアが高くなるのではなく、形式に不慣れなため十分に実力を発揮できないことが考えられるからです。特に日本人の場合、普段学校等で慣れ親しんできた英語のテストとは、問題量やスピードがかなり違っているため、初めて受験した人はとまどってしまい、結果的に実力を発揮できないまま終わってしまったという人も見受けられます。そのため、自分の実力を正確に測るためにも形式には事前に慣れておくことが望ましいのです。最後の「受験指導による得点効果」は予想問題が当ることによる得点上の有利さを意味します。しかしこれは、ある程度出題範囲が決まっている場合には可能ですが、TOEIC は出題範囲が限定されておらず、いにかえればすべての範囲から出題されるものです。ですから、同じ問題を予想することは不可能ですし、いくつか同じ単語や熟語、構文といったものを当てることができたとしても、全体のスコアにあたる影響は微々たるものです。つまり、受験指導による得点効果にたいしても対応できていることとなります。TOEIC はあらゆる点で、実力以外の要因によって評価基準にゆがみが生じないように、細心の注意を払って設計されています。付け焼き刃の受験対策は TOEIC には通用しないにご理解ください。但し、英会話学校や市販の教材によって TOEIC の類例を多数練習することは、決して無駄ではありません。TOEIC と同じようなスピードで英語を聞き、速いスピードで英語を読む練習を続けることにより、スピード対応能力がつき結果として TOEIC スコアがあがれば、それはまさに英語の実力がついたこと⁴ を示している のです。

See! <http://www.toEIC.or.jp/toEIC/faq/main.html>

² 上記の強調は井上の一存。アンケートで一体何がどう測られているのか? というのを見極める事は決して易しくない! 調査する側はもっと命がけで! そして、より良い調査方式を参考意見として学生自治会へ手渡そう

³ 上記の強調は井上の責任!

⁴ アンケートを取る事自身ではなく、「その講義の授業の質をどうしたら高める事ができるか」が目的であった事をお忘れなく!!