

Trong tiến trình xây dựng chất lượng giáo dục, một công tác cốt lõi, có tác động mang tính hệ thống đến chất lượng là công tác đánh giá người học hay hiểu một cách đơn giản là việc ra đề thi và chấm điểm.

Thoạt trông thì đơn giản, thế nhưng nhiều người kể cả các giáo viên đôi khi vẫn chưa nhận thức rõ cơ sở lý luận, tầm quan trọng và mức độ ảnh hưởng của công việc này trong chuỗi giá trị hình thành chất lượng giáo dục.

Một vài quan sát

Trước khi đi vào vấn đề này, hãy cùng quan sát một số hiện tượng thường gặp trong đời sống xã hội (gọi là đời sống xã hội vì giáo dục thực chất là câu chuyện của mọi nhà). Hiện tượng thứ nhất khá phổ biến trong các trường đại học là chuyện các “Dũng sĩ diệt sinh viên”. Sinh viên các trường luôn truyền khẩu về một vài thầy cô luôn rất nghiêm khắc trong việc ra đề và chấm điểm. Có những môn của một thầy/ cô, năm nào, khóa nào thi lần đầu cũng chỉ 20-30% là được điểm trên trung bình (cao nhất có khi chỉ 6-7 điểm/ 10), còn 70-80% thi lại. Có vấn đề gì ở đây? Môn học quá khó? Sinh viên quá kém? hay phương pháp giảng dạy và đánh giá có vấn đề? Tương tự, hiện tượng thứ hai trước đây vài năm cũng rất phổ biến— đó là việc các đề thi vào đại học, đặc biệt là các đại học nổi tiếng, cực kỳ hóc búa, có năm hầu hết thí sinh đều dưới trung bình. Có những đề thi mà thậm chí cả những thầy dạy luyện thi cũng không giải nổi. Cũng gần đó câu hỏi cần được đặt ra. Học sinh quá kém hay phương pháp đánh giá có vấn đề? Hiện tượng thứ ba ít được chú ý hơn nhưng cũng không kém phần phổ biến, và nghiêm trọng. Đó là việc có rất nhiều

PHƯƠNG PHÁP ĐÁNH GIÁ Độ chính xác và độ tin cậy

VŨ THẾ DŨNG

môn học ở bậc phổ thông cũng như đại học giáo viên hầu như không cho điểm tối đa 10/10 thậm chí 9/10 mà chỉ cho điểm cao nhất là 8/10. Hình như có một thứ quan điểm ngầm hiểu là không ai có thể đạt được điểm tối đa vì kiến thức là vô hạn và chấm 8 điểm là để học sinh hiểu rằng mình còn chưa hoàn thiện còn cần phải liên tục cố gắng. Một số thầy cô ở bậc đại học còn chấm điểm theo phương pháp điểm trừ. Tức là chấm điểm bằng cách trừ dần vào các lỗi sinh viên mắc phải trong bài thi, chứ không chỉ chấm điểm phần làm đúng. Và theo phương pháp này có nhiều sinh viên còn bị điểm âm. Cách ra đề, chấm điểm hiện nay tùy thuộc vào từng giáo viên và dẫn đến hệ quả là điểm số không được chuẩn hóa hay điểm của các khóa, các lớp không thể so sánh với nhau. 6 điểm của thầy A thì không thể so sánh với 6 điểm của cô B vì mỗi người có quan điểm cho điểm khác nhau. Chuyện này không chỉ bất công với người học mà còn làm cho hệ thống điểm số không thể dùng được trong so sánh chất lượng giáo dục, nên rất khó tiến hành các nghiên cứu định lượng về chất lượng giáo dục. Trước đây, vấn đề điểm số không quá quan trọng thế nhưng khi học sinh, sinh viên bắt đầu ra nước ngoài du học, họ phải nộp bằng điểm cho các trường nước ngoài thì mới thấy có vấn đề lớn. Nhiều sinh viên tốt nghiệp là thủ khoa các trường nhưng điểm trung bình chỉ khoảng 8/10. Trong khi các thủ khoa nước ngoài thường là toàn A (full A) tức là luôn đạt điểm tối đa. Nếu so một cách cơ học thì 8/10 chỉ cỡ B hay

B+ ở nước ngoài. Dẫn đến thiệt thòi và giảm cơ hội cho du học sinh.

Cần cách tiếp cận khác

Các hiện tượng và vấn đề trên có lẽ đều có cùng một cái gốc là quan điểm và phương pháp lạc hậu trong đánh giá người học. Có một số quan điểm mới trong đánh giá. Thứ nhất, đánh giá người học phải chú trọng đánh giá quá trình chứ không đánh giá bằng một bài thi duy nhất. Sự chuyển đổi này là do đánh giá quá trình tăng tính chính xác của việc đánh giá. Đánh giá bằng một bài thi duy nhất cuối khóa sẽ rất rủi ro và có thể phiến diện. Thứ hai, chấm điểm tương đối thay vì chấm điểm tuyệt đối. Chấm điểm tuyệt đối chỉ căn cứ vào việc người học làm được chính xác bao nhiêu phần trăm của đề thi, còn chấm điểm tương đối ngoài việc so sánh với chuẩn tuyệt đối còn so sánh tương quan với các việc thực hiện bài thi của các thành viên khác. Chẳng hạn, một bài thi quá khó dẫn đến sinh viên giỏi nhất cũng chỉ hoàn thành được 80% - như vậy theo phương pháp tuyệt đối sinh viên này được 8 điểm. Phương pháp tương đối cho rằng 80% là mốc cao nhất sinh viên thực hiện được và do vậy người được 80% vẫn có thể đạt điểm tối đa là 10 và điểm của các sinh viên khác được tính tương đối theo điểm của người đứng đầu.

Thông thường một đề thi tốt phải phân loại được trình độ của người học. Phân phối điểm của một lớp hay một khóa phải cân đối giữa điểm cao, thấp và trung bình. Quá

lịch về bên nào cũng đều cần đánh giá lại tính chính xác của đề thi. Cũng cần thấy rằng điểm 10/10 thể hiện việc sinh viên hoàn thành chính xác các nội dung đã được học. Nó không làm cho người học trở nên kiêu ngạo vì thành tích nếu điểm số được hiểu và truyền thông chính xác. Cho nên rất cần cho điểm tối đa để khuyến khích người học. Điểm này dẫn đến ý thứ ba. Cần đánh giá thông qua khuyến khích các thành tích chứ không chú trọng vào phạt các lỗi lầm (trừ điểm các lỗi). Thoạt trông thì giống nhau, nhưng thực ra hai cách tiếp cận này dẫn đến các hệ quả khác nhau. Thường cho thành tích khuyến khích người học cố gắng tìm tòi và học hỏi những cái mới, trong khi phạt khi phạm lỗi khuyến khích người học chú ý sửa lỗi, không mắc lỗi nhưng lại hạn chế tìm tòi, sáng tạo vì tìm tòi, sáng tạo cái mới rất dễ phạm lỗi, dẫn đến bị phạt.

Độ chính xác và độ tin cậy

Một đề thi tốt được xem là một thang đo tốt. Thang đo này phải đạt được hai yêu cầu bắt buộc: độ chính xác (validity) và độ tin cậy (reliability). Độ chính xác là việc bài thi đo được cái cần đo. Chẳng hạn một đề thi cuối khóa là nhằm mục tiêu đo được việc tiếp thu những kiến thức và kỹ năng đã được truyền đạt trong quá trình giảng dạy chứ không phải để đánh đố hay ra những câu hỏi, nội dung không được truyền đạt trong quá trình dạy và học. Những bài thi đánh đố thông thường là những bài thi không đạt yêu cầu về độ chính xác vì nó không được thiết kế để đo cái cần đo.

Có hai loại độ chính xác thường đề cập trong các phương pháp đánh giá. Đầu tiên là độ chính xác về nội dung (content validity). Đây là yêu cầu thiết thân của mọi đề thi hay thang đo. Bài thi đạt độ chính xác về nội dung khi nó bao quát được toàn bộ nội dung đã giảng dạy. Để làm điều này phải xác định được chính

xác phạm vi và nội dung giảng dạy của một môn hay một chương trình (cho các kỳ thi tốt nghiệp hay vào đại học). Trên cơ sở đó thực hiện việc lấy mẫu các câu hỏi có tính chất đại diện cho toàn bộ nội dung. Lúc này các nguyên lý của thống kê được áp dụng. Các câu hỏi phải đại diện được cho các nhóm nội dung và có phân bố đồng đều về độ khó. Rất nhiều trường hợp, bài thi chỉ hỏi một phần rất nhỏ của nội dung giảng dạy dẫn đến học tử và rủi ro cho sinh viên. Hoặc hỏi toàn những câu quá khó không nằm trong nội dung giảng dạy – lúc này độ chính xác về nội dung của thang đo là kém. Điểm không phản ánh chất lượng dạy và học, không đóng vai trò định hướng và thang đo chất lượng.

Loại độ chính xác thứ hai, có yêu cầu cao hơn là độ chính xác trên phương diện dự báo (predictive validity or criterion validity). Mỗi môn học hay chương trình học đều có mục tiêu của nó, do vậy các bài thi cũng phải được thiết kế để đo được việc thực hiện các mục tiêu đó. Chẳng hạn môn học thống kê trong kinh doanh có nhiều mục tiêu nhưng mục tiêu cụ thể đầu tiên là để phục vụ việc làm tốt công tác phân tích số liệu trong các môn học có liên quan sau đó như Nghiên cứu tiếp thị. Vậy một bài thi thống kê được đánh giá đạt độ chính xác về dự báo khi sinh viên đạt điểm cao trong bài thi đó tiếp tục làm tốt công việc phân tích số liệu được giao trong các môn học kế tiếp. Hay sinh viên vượt qua kỳ tốt nghiệp phải làm tốt công việc anh ta được thiết kế để làm. Và bài thi vào đại học phải chọn được những người có xác suất thành công cao trong trường đại học và loại những người có khả năng thành công thấp hơn. Nếu không phản ánh được tiêu chí dự báo, bài thi không chính xác. Tất nhiên ở điểm này cũng có thể phải quay câu hỏi về phía chương trình đào tạo đã đạt độ chính xác chưa, hay nó đã được thiết kế chính

xác để những ai hoàn thành nó có thể thực hiện tốt các công việc mà nó thiết kế để làm chưa? Phạm trù độ chính xác do vậy hoàn toàn có thể áp dụng cho việc đánh giá các chương trình đào tạo chứ không chỉ giới hạn đánh giá các bài thi.

Để thực hiện việc đánh giá độ chính xác trên phương diện dự báo của các đề thi quốc gia hiện nay, chúng ta có thể thực hiện các nghiên cứu đơn giản như tính các cặp hệ số tương quan (correlation) giữa điểm thi đại học, điểm thi tốt nghiệp trung học, điểm trung bình trung học với điểm trung bình năm thứ nhất, năm thứ hai, và điểm tốt nghiệp đại học; rồi với điểm học đại học với thành công trong xin việc làm trong 3 tháng đầu tiên sau khi tốt nghiệp... Hệ số tương quan giữa các cặp số liệu càng cao thì độ chính xác càng tốt, và ngược lại. Hiếm có thang đo hay bài thi nào hoàn toàn chính xác trên mọi phương diện, do vậy để tăng độ chính xác người ta phải đa dạng hóa các thang đo, chuyển từ một bài thi sang việc sử dụng nhiều tiêu chí khác nhau như điểm thi tốt nghiệp, điểm thường cho học sinh giỏi v.v.v. Vấn đề đang tranh luận hiện nay về điểm thường hay miễn thi đại học cho học sinh đạt giải toàn quốc lại không sai ở triết lý đánh giá (rất cần đa dạng hóa), mà sai ở cách thức triển khai và thực hiện.

Sau khi đạt được độ chính xác, thang đo phải đạt được độ tin cậy hay nói cách khác nó phải cho các kết quả như nhau trong các lần đo khác nhau. Có thể nhìn yêu cầu này dưới các góc độ: (1) sinh viên cùng 1 trình độ sẽ đạt cùng 1 thang điểm trong 1 bài thi, (2) 1 sinh viên sẽ đạt cùng 1 thang điểm trong các bài thi cùng trình độ, (3) điểm số có thể so sánh được qua các thời gian khác nhau. Đây là yêu cầu chuẩn hóa bài thi. Chính vì vậy thi trắc nghiệm trở nên có ưu thế hơn so với hình thức tự luận vì dễ chuẩn hóa (đảm bảo độ tin cậy) và nhiều câu hỏi nên có thể đi vào nhiều nội dung khác nhau (đảm bảo độ chính xác về nội

dung).

Chúng ta cần nghiêm túc học tập kinh nghiệm của các nước trong vấn đề này và cần đưa khoa học đánh giá, đo lường vào trong giáo dục. Chẳng hạn, bài thi SAT – bài thi tuyển đầu vào được sử dụng bởi hơn 80% các trường đại học của Mỹ – đã được xây dựng từ năm 1926, đến nay đã có hơn 60 triệu sinh viên trải qua kỳ thi này. Để có một bài thi chuẩn hóa với độ chính xác cao, qui trình xây dựng bài thi hết sức khoa học và chặt chẽ. Nó trải qua 12 bước:

1. Xác định phạm vi nội dung ra đề.
2. Viết câu hỏi: do các giáo viên, các nhà khoa học phụ trách.
3. Đánh giá lại câu hỏi về mặt nội dung.
4. Đánh giá lại câu hỏi về tính công bằng (fairness).
5. Kiểm tra thử câu hỏi với học sinh.
6. Phân tích thống kê từ kết quả kiểm tra thử: đánh giá độ khó của câu hỏi, đánh giá khả năng phân biệt trình độ học sinh của câu hỏi, đánh giá kết quả của các học sinh cùng trình độ nhưng thuộc các thành phần khác nhau (tôn giáo, dân tộc, giới tính) để xem xét tính công bằng.
7. Hiệu chỉnh và hình thành bản nháp của bài thi.
8. Thẩm định của các chuyên gia bên ngoài.
9. Thẩm định của hội đồng chuyên môn: bao gồm đại diện của cố vấn của các trường trung học, chuyên gia tuyển sinh của các trường đại học, các nhà giáo dục, các giáo viên.
10. Biên tập, sửa bản in, in, và phân phối.
11. Tổ chức thi.
12. Phân tích thống kê sau khi thi: thực hiện phân tích giống như đã thực hiện ở bước 6 để đảm bảo kết quả thi thử và thi thật không chênh.

Ngoài ra ETS (tổ chức xây dựng và phát triển các bài thi chuẩn hóa bao gồm TOEFL, SAT, GMAT, GRE...) còn thường xuyên thực hiện

(Tiếp theo trang 8)

Đề xuất kết nối ...

- Thành lập một cơ chế giám sát công bố thông tin chéo giữa các thị trường tham gia vào hoạt động hợp tác chống vi phạm cần được thiết lập. Đồng thời, thành lập hệ thống giám sát chéo giữa các thành viên thị trường như CTCK, TTLKCK, ngân hàng chỉ định thanh toán...

Sự kiện kết nối TTCK VN sẽ đánh dấu bước phát triển phi mã của TTCK VN, tạo sân chơi bình đẳng cho các NĐT trên TTCK các nước khu vực, thực hiện cam kết hội nhập thị trường tài chính VN, hội nhập kinh tế quốc tế góp phần đẩy mạnh tốc độ tăng trưởng kinh tế VN ■

Ghi chú:

[1] Báo cáo thường niên 2005, TTGDCK TP.HCM, trang 17

Tài liệu tham khảo:

1. Quyết định số 163/2003/QĐ-TTg, ngày 05/08/2003 của Thủ tướng Chính phủ phê duyệt chiến lược phát triển TTCK VN đến năm 2010.
2. Quyết định số 898/QĐ-BTC, ngày 20.2.2006 của Bộ trưởng Bộ tài chính ban hành Kế hoạch phát triển thị trường chứng khoán VN 2006-2010.
3. Báo cáo thường niên Trung tâm giao dịch chứng khoán TP.HCM, năm 2005
4. Tài liệu của UOB ASIA LIMITED
5. Website worldbank.org.vn; euronet.com; sgx.com; vse.org.vn; ssc.gov.vn

(Tiếp theo trang 12)

Hướng phát triển cho thị trường OTC VN

Thứ ba là tăng cường hệ thống giám sát, qui định rõ trong Luật chứng khoán: chức năng nhiệm vụ, quyền hạn của nhà quản lý thị trường OTC trong công tác giám sát thị trường. Phải có sự phối hợp giám sát giữa UBCKNN và tổ chức tự quản của thị trường OTC. Phân định chức năng nhiệm vụ của ủy ban hay tổ chức tự quản trong việc xử lý vi phạm trên thị trường OTC.

Thứ tư là phát triển hệ thống công bố thông tin hiện đại, đảm bảo thông tin không có nội gián và thể hiện được tính công bằng trong việc tiếp nhận thông tin của các nhà đầu tư. Đặc biệt trên thị trường OTC thời gian qua việc công bố thông tin của các công ty chưa đủ tiêu chuẩn niêm yết chưa đến được các nhà đầu tư một cách trung thực, cập nhật và phổ biến. UBCKNN phải có những qui định cụ thể và nghiêm ngặt đối với những công ty này trong việc cam kết công bố thông tin ra thị trường. Ngoài các công ty bán chứng khoán phải cung cấp thông tin về tình hình tài chính công ty, thì các công ty làm môi giới cũng phải có nghĩa vụ cung cấp thông tin đầy đủ cho nhà đầu tư về tình hình của thị trường như số lượng, giá cả chứng khoán mua bán.

Thị trường OTC phát triển sẽ bổ sung vào hoạt động giao dịch chứng khoán thứ cấp, tạo động lực cho sự phát triển của thị trường sơ cấp qua việc tạo tính thanh khoản cho chứng khoán của các doanh nghiệp vừa và nhỏ. Thị trường này cũng làm phong phú hơn hoạt động huy động vốn và đầu tư trên một TTCK còn non trẻ, góp phần làm hoàn thiện TTCK Việt Nam ■

Tài liệu tham khảo:

1. TS. Trần Cao Nguyên, Đề tài nghiên cứu khoa học cấp Bộ "Chiến lược phát triển TTCK Việt Nam đến năm 2010", Hà Nội 2000.
2. Quyết định 163/2003/QĐ-TTg ngày 5/8/2003 phê duyệt chiến lược phát triển TTCK Việt Nam đến năm 2010.
3. "Chứng khoán và TTCK, những kiến thức cơ bản", Ủy ban chứng khoán Nhà nước.
4. ThS Phạm Nguyễn Hoàng, Thị trường chứng khoán OTC, Tạp chí Chứng khoán số 4.2006.

các nghiên cứu để đánh giá độ chính xác và tin cậy của các bài thi của mình, như thực hiện các phân tích hệ số tương quan giữa điểm thi SAT với

điểm trung bình tích lũy trong năm đầu tiên học đại học. Hệ số này luôn rất cao thể hiện tính chính xác của đề thi ■