



PHÂN TÍCH CÂU HỎI TỰ LUẬN TRONG ĐÁNH GIÁ KẾT QUẢ HỌC TẬP CỦA HỌC SINH MÔN TOÁN TRUNG HỌC PHỔ THÔNG

ThS. HÀ XUÂN THÀNH

Cục Khảo thí và Kiểm định chất lượng giáo dục

Đánh giá kết quả học tập là một khâu quan trọng trong quá trình dạy học. Trong những năm trở lại đây, việc sử dụng kết hợp câu hỏi trắc nghiệm khách quan (TNKQ) và tự luận trong các đề kiểm tra đã trở nên phổ biến trong đánh giá kết quả học tập các môn học nói chung và môn Toán trung học phổ thông (THPT) nói riêng. Tuy nhiên, có thể thấy rằng, còn nhiều ý kiến khác nhau về việc sử dụng loại hình câu hỏi nào cho các đánh giá như vậy, đặc biệt là trong các đánh giá quan trọng như kì thi tốt nghiệp THPT, kì thi tuyển sinh đại học, cao đẳng. Nhiều ý kiến cho rằng nên sử dụng các câu hỏi TNKQ vì chúng bao phủ được hết phạm vi chương trình và kết quả chấm mang tính khách quan hơn. Một số ý kiến khác cho rằng, sử dụng câu hỏi tự luận mới có thể đánh giá được các mức độ tư duy cao cũng như khả năng diễn đạt của học sinh. Do đó, vấn đề sử dụng loại hình câu hỏi nào cho đánh giá cũng đang là vấn đề cần được quan tâm tìm hiểu. Tuy nhiên, cho dù bất kì một loại hình câu hỏi nào được sử dụng thì điều quan trọng hơn cả là mỗi câu hỏi cần đảm bảo có chất lượng theo một nghĩa rộng nhất. Điều này có nghĩa, bản thân câu hỏi đó là có chất lượng và nó thể hiện được vai trò khi đưa vào một đề kiểm tra nào đó. Một trong những yêu cầu đánh giá của thế giới hiện nay là xu thế đánh giá năng lực, trong đó, có năng lực vận dụng những điều đã học vào các tình huống thực tiễn, mà các tình huống thực tiễn được phát triển vận dụng thường thể hiện dưới những bài toán tự luận. Trong bài viết này, dựa trên cơ sở của lí thuyết đánh giá và đo lường hiện đại, chúng tôi cố gắng đưa ra một số vấn đề phân tích, đánh giá câu hỏi tự luận nhằm nâng cao chất lượng câu hỏi tự luận trong việc đánh giá kết quả học tập môn Toán của học sinh THPT.

1. Tổng quan về câu hỏi tự luận

Trên thế giới, người ta chia thành hai loại hình câu hỏi chủ yếu, đó là câu hỏi mang tính lựa chọn (selected-response item) như câu hỏi có nhiều phương án lựa chọn, câu hỏi đúng sai, điền khuyết, trả lời ngắn hoặc ghép đôi và các câu hỏi mang tính xây dựng câu trả lời (constructed-response item). Loại hình câu hỏi này được biết đến như các bài tự luận (essay) [5].

Có nhiều quan niệm khác nhau về một câu hỏi tự luận, tuy nhiên bài tự luận trong đánh giá kết quả học tập có một số đặc trưng như:

- (i) Đưa ra các yêu cầu cụ thể về nhiệm vụ học sinh cần phải thực hiện;
- (ii) Học sinh tự đưa ra câu trả lời để giải quyết nhiệm vụ đó trên cơ sở các kiến thức, kĩ năng, thái độ đã có;
- (iii) Việc chấm điểm được thực hiện bằng tay theo một đáp án và biểu điểm cho trước. Thông thường, câu hỏi tự luận được biên soạn để đánh giá kết quả học tập của học sinh ở các cấp độ tư duy cao như hiểu, phân tích, tổng hợp, đánh giá. Để viết được một câu hỏi tự luận không khó nhưng một câu hỏi tự luận có chất lượng, đánh giá đúng mục tiêu đề ra là một vấn đề không phải dễ dàng giải quyết được. Nhiều tác giả đã đưa ra các quy tắc để viết câu hỏi tự luận có chất lượng, trong đó, một nguyên tắc xuyên suốt là câu hỏi phải được đánh giá, phân tích một cách cẩn thận trước khi sử dụng chính thức.

Câu hỏi tự luận thể hiện được một số ưu điểm như dễ biên soạn câu hỏi, tốn ít thời gian và có thể dựa nhiều vào kinh nghiệm, có thể cung cấp thông tin cho việc đánh giá kĩ năng diễn đạt, khả năng suy luận lôgic và tạo điều kiện để học sinh thể hiện chính kiến của mình khi giải thích, có thể đánh giá được năng lực sáng tạo của học viên ở mức độ cao. Tuy nhiên, bên cạnh đó, nó cũng có một số nhược điểm như chỉ có thể kiểm tra được một số lĩnh vực cơ bản trong một khoảng thời gian nhất định nên dễ tạo cho họ có thói quen học tủ, kết quả chấm điểm dễ bị ảnh hưởng bởi quan niệm và thái độ, sức khoẻ người chấm và không kiểm tra được phạm vi kiến thức rộng.

2. Phân tích, đánh giá câu hỏi tự luận

Bất kì một câu hỏi nào khi được biên soạn đều cần được xem xét chất lượng một cách kĩ lưỡng. Việc xem xét chất lượng câu hỏi thông qua việc phân tích thống kê đã là một xu thế trên thế giới. Tùy vào mục đích đánh giá mà quyết định mức độ của việc phân tích câu hỏi. Đối với các đánh giá thường xuyên trên lớp học thì câu hỏi tự luận có thể được giáo viên biên soạn dựa vào kinh nghiệm giảng dạy của cá nhân, tuy nhiên trong các đánh giá quan trọng, chẳng hạn như

khảo sát chất lượng học tập hoặc các kì thi thì việc phân tích câu hỏi tự luận đóng một vai trò rất quan trọng và là một khâu không thể thiếu trong quá trình thiết kế đánh giá đó. Việc phân tích thường tiến hành theo hai phương pháp. Phương pháp chuyên gia nhằm lấy ý kiến của các chuyên gia môn học, các giáo viên giỏi về chất lượng câu hỏi dựa vào kinh nghiệm cá nhân thông qua một hệ thống các tiêu chí. Phương pháp thứ hai là sử dụng các phân tích thống kê để đánh giá chất lượng.

2.1. Một số tiêu chí xem xét chất lượng của câu hỏi tự luận

Theo Anthony J. Nitko [4], để đánh giá câu hỏi tự luận, một số tiêu chí chủ yếu cần được xem xét là:

1. Phản ánh đúng mục tiêu đánh giá hay không và có phù hợp với ma trận để kiểm tra
2. Đòi hỏi học sinh phải vận dụng kiến thức vào các tình huống mới
3. Yêu cầu của câu hỏi phù hợp với trình độ và nhận thức của học sinh
4. Để đạt được điểm cao, học sinh có phải chứng minh quan điểm của mình hơn là nhận biết về thực tế
5. Ngôn ngữ sử dụng trong câu hỏi truyền tải được hết những yêu cầu của cán bộ ra đề đến học sinh.

Ngoài 5 tiêu chí cơ bản nói trên, câu hỏi tự luận còn kèm theo hướng dẫn chấm điểm. Thiết kế hướng dẫn chấm điểm là một công việc phức tạp, đòi hỏi nhiều kĩ thuật, việc trình bày chi tiết sẽ được giới thiệu trong một dịp khác.

2.2. Đánh giá thống kê câu hỏi tự luận

Sự phát triển mạnh mẽ của lí thuyết đánh giá và đo lường đã tạo nên một số thay đổi lớn trong việc phân tích câu hỏi. Mỗi câu hỏi sau khi phân tích sẽ có các tham số đặc trưng như độ khó, độ phân biệt, độ phỏng đoán. Các tham số này sẽ là một trong các cơ sở quan trọng để xác định chất lượng câu hỏi. Với câu hỏi TNKQ, việc phân tích thống kê sử dụng các thành tựu của lí thuyết trắc nghiệm hiện đại đã trở nên rất phổ biến trên thế giới và cũng khá phổ biến ở nước ta trong những năm gần đây. Tuy nhiên, việc phân tích như vậy đối với câu hỏi tự luận ở nước ta chưa được quan tâm thích đáng nhằm nâng cao chất lượng câu hỏi tự luận.

Để minh họa cho việc đánh giá thống kê các câu hỏi tự luận, chúng tôi tiến hành phân tích đề thi tốt nghiệp môn Toán THPT năm 2010 căn cứ vào kết quả làm bài của 222 học sinh lớp 12. Phần mềm CONQUEST được sử dụng để phân tích các số liệu thu được.

Để thực hiện việc phân tích các câu hỏi, trước hết chúng tôi sử dụng Hướng dẫn chấm điểm của Bộ Giáo dục và Đào tạo cho từng câu hỏi. Chẳng hạn, xét câu

hỏi 2c trong đề.

Cho hàm số $f(x) = x - 2\sqrt{x^2 + 12}$. Giải bất phương trình $f'(x) \leq 0$.

Hướng dẫn chấm của câu hỏi này như sau:

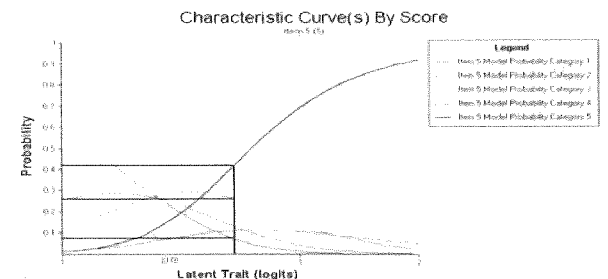
Trên tập xác định $D = \mathbb{R}$ của hàm số $f(x)$, ta có: $f'(x) = 1 - \frac{2x}{\sqrt{x^2 + 12}}$	0,25
Do đó: $f'(x) \leq 0 \Leftrightarrow \sqrt{x^2 + 12} \leq 2x$	0,25
$\Leftrightarrow \begin{cases} x \geq 0 \\ x^2 \geq 4 \end{cases}$	0,25
$\Leftrightarrow x \geq 2.$	0,25

Với hướng dẫn chấm như vậy, việc mã hoá sẽ được tiến hành bằng việc quy đổi 0.25 điểm thành một đơn vị tính điểm, nghĩa là câu hỏi 1 điểm này sẽ được quy đổi thành câu hỏi có 4 đơn vị điểm. Ví dụ, một học sinh được được 0.5 điểm tương ứng với 2 đơn vị điểm, một học sinh được 0.75 điểm tương ứng với 3 đơn vị điểm, học sinh không làm câu hỏi này gán mã 9. Các mã hoá như vậy được sử dụng để lập file dữ liệu phân tích. Kết quả chạy phần mềm CONQUEST đối với câu hỏi này cho ta một số kết quả chủ yếu sau:

```
item:5 (5)
Cases for this item 222 Discrimination 0.69
Item Threshold(s): -0.75 -0.13 0.39 0.59 Weighted
MNSQ 1.00
Item Delta(s): -0.19 -0.24 1.29 -0.86
-----
Label Score Count % of tot Pt Bis t (p) PV1Avg:1
PV1 SD:1
```

```
0 0.00 44 19.82 -0.38 -6.08(.000) 0.12 0.29
1 1.00 32 14.41 -0.20 -3.03(.003) 0.21 0.33
2 2.00 53 23.87 0.09 1.40(.162) 0.48 0.36
3 3.00 22 9.91 0.08 1.26(.207) 0.45 0.36
4 4.00 60 27.03 0.54 9.62(.000) 0.88 0.40
9 0.00 11 4.95 -0.39 -6.37(.000) -0.14 0.46
```

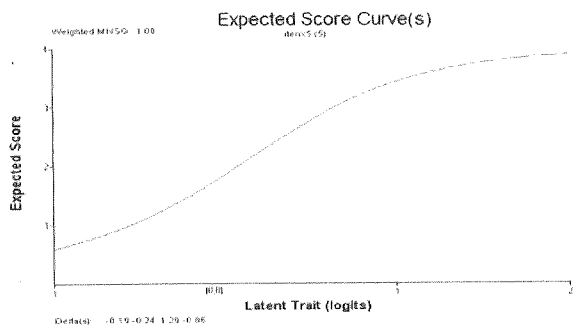
Đường cong đặc trưng và đường cong thông tin của câu hỏi được biểu thị trong các hình dưới đây:



Hình 1. Đường cong đặc trưng câu hỏi theo điểm số

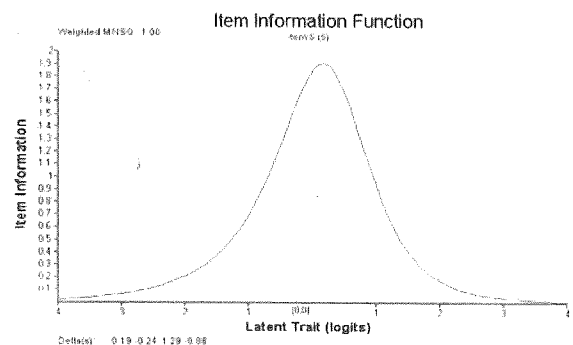
Đường cong đặc trưng câu hỏi theo điểm số cung cấp cho ta mối quan hệ giữa xác suất trả lời đúng của

mỗi thí sinh có năng lực θ ứng với mỗi điểm số tương ứng. Chẳng hạn, ứng với năng lực $\theta = 0.45$, xác suất để học sinh này được 0.25 điểm xấp xỉ là 0.08, xác suất để được 0.75 điểm xấp xỉ là 0.275,...



Hình 2. Đường cong đặc trưng kì vọng của câu hỏi

Đường cong đặc trưng kì vọng cho thấy mối quan hệ giữa năng lực của thí sinh với tổng điểm mong đợi đạt được đối với câu hỏi. Đây cũng là cơ sở để xác định một câu hỏi là dễ hay khó đối với một nhóm thí sinh nào đó.



Hình 3. Đường cong thông tin câu hỏi

Các thông tin thống kê như bảng ở trên cùng với đường cong thông tin và đường cong đặc trưng sẽ giúp cho việc đánh giá tốt nhất có thể chất lượng của câu hỏi. Nhìn vào kết quả trên và đường cong đặc trưng kì vọng của câu hỏi, ta thấy đây là câu hỏi có độ khó trung bình, có độ phân biệt rất cao ($a = 0.69$), thông tin thu được nhiều tập trung ở mức năng lực trung bình thể hiện qua đường cong thông tin ở Hình 3.

Bên cạnh các thông tin câu hỏi, thông qua phần mềm CONQUEST thì năng lực các học sinh tham gia làm bài kiểm tra cũng được ước lượng. Nếu thỏa mãn một số điều kiện nhất định thì các ước tính năng lực này có thể được sử dụng như điểm năng lực của thí sinh về lĩnh vực đang đánh giá. Đây là một cơ sở quan trọng trong việc đánh giá theo hướng tiếp cận năng lực.

Trên đây chỉ là một số minh họa ban đầu trong việc phân tích câu hỏi tự luận mà không đi sâu vào một số kĩ thuật chi tiết khác. Tuy nhiên, thông qua các ví dụ trên, có thể thấy việc phân tích câu hỏi tự luận sẽ mang lại các thông tin hữu ích nhằm cải thiện chất lượng câu hỏi trước khi sử dụng. Việc phân tích có thể tiến hành thông qua kinh nghiệm cũng như có thể phân tích thống kê với việc ứng dụng một số phần mềm dựa trên lí thuyết đánh giá và đo lường. Nếu thiết kế được câu hỏi tự luận có chất lượng để thiết kế để kiểm tra sẽ góp phần nâng cao chất lượng công tác đánh giá, từ đó cung cấp các thông tin cần thiết phục vụ cho quá trình dạy học.

TÀI LIỆU THAM KHẢO

1. Đặng Xuân Cương, *Sử dụng thống kê phân tích trong việc nâng cao chất lượng câu hỏi kiểm tra phục vụ công tác đánh giá kết quả học tập môn Toán của học sinh THCS*, Kỷ yếu Hội thảo toàn quốc "Xác suất – Thống kê: Nghiên cứu, Ứng dụng và Giảng dạy", Viện Toán học Việt Nam, 2010.
2. Trần Kiều, *Một số vấn đề chung về đánh giá chất lượng giáo dục phổ thông (Tài liệu đánh giá cho các lớp tập huấn của Dự án Phát triển Giáo dục THCS II)*, 2006.
3. Lâm Quang Thiệp, *Đo lường trong giáo dục, Lí thuyết và thực hành*, Nhà xuất bản Đại học Quốc gia Hà Nội, 2010.
4. Susan M. Brookhart, Anthony J. Nitko, *Education Assessment of students*, Pearson Education, Inc, Upper Saddle River, New Jersey, 2007.
5. Susan M. Brookhart, Anthony J. Nitko, *Assessment and Grading in classroom*, Pearson Education, Inc, Upper Saddle River, New Jersey, 2008.
6. Margret L. Wu, *Acer ConQuest, Version 2.0: Generalised Item Response Modelling software*. In Raymond J. Adams, Mark R. Wilson, Samuel A.Haldane (ed.). Australian Council for Educational Research, 2007.

SUMMARY

The author goes into the analysis of essay questions in upper secondary school mathematics (quality checking indicators, statistical review) for quality improvement of questions, test design with essay questions as basis for students assessment, providing information for teaching.