

## ĐÁNH GIÁ KẾT QUẢ HỌC TẬP MÔN HỌC HOÁ ĐẠI CƯƠNG BẰNG HÌNH THỨC TRẮC NGHIỆM KHÁCH QUAN NHIỀU LỰA CHỌN

VÕ THỊ HỒNG TỊNH<sup>1</sup>

### 1. Mở đầu

Quá trình đánh giá thành quả học tập của sinh viên ở giáo dục bậc cao - Đại học và Cao đẳng - hiện nay ở nước ta còn nhiều khiếm khuyết, thể hiện qua:

- Sự thiếu kiến thức về phương pháp đánh giá của một số không nhỏ giáo viên.
- Sử dụng quá trình đánh giá sai mục đích chỉ chú trọng đến chức năng tuyển chọn, phân loại sinh viên trong khi coi nhẹ chức năng hỗ trợ cho quá trình dạy và học.

Hình thức đánh giá xưa nay vẫn dùng quen thuộc trong các trường đại học thuộc các ngành khoa học tự nhiên là những bài kiểm tra viết bao gồm những câu hỏi lí thuyết và những bài tập lí giải hay tính toán. Đây là phương tiện khảo sát thành quả học tập hữu hiệu và cần thiết, tuy nhiên nó cũng có những mặt hạn chế sau:

- 1) Khi cần khảo sát thành quả học tập của một số đông học sinh.
- 2) Khi muốn có những điểm số đáng tin cậy, không phụ thuộc vào chủ quan của người chấm bài.
- 3) Khi việc thi và kiểm tra đòi hỏi các yếu tố công bằng, vô tư và chính xác
- 4) Khi muốn ngăn ngừa nạn học tủ, học vẹt và gian lận trong thi cử.

Vấn đề đặt ra là cần có những hình thức thi và kiểm tra đa dạng song song với hình thức thi viết truyền thống. Hình thức trắc nghiệm khách quan (TNKQ) nhiều lựa chọn đã được đưa vào sử dụng một cách phổ biến trong việc đánh giá kết quả học tập của sinh viên đối với các bộ môn ngoại ngữ, còn ở các bộ môn khoa học tự nhiên thì hiếm gặp, nhất là đối với bộ môn như Hóa Đại cương.

Việc xây dựng và tổ chức áp dụng TNKQ ở các nước tiên tiến đã đi vào nề nếp từ nhiều năm. Ở Việt Nam, nhìn chung hình thức trắc nghiệm khách quan áp dụng trong thi kiểm tra kết quả học tập của học sinh chưa mấy phổ biến. Ở các bộ

---

<sup>1</sup> Thạc sĩ, Giảng viên Khoa Hoá, Trường Đại học Sư phạm TP.HCM.

môn Hoá Đại cương của các trường đại học tại Thành phố Hồ Chí Minh cũng đã bắt đầu sử dụng trắc nghiệm nhiều lựa chọn trong các kì thi kết thúc môn học, nhưng bộ đề các câu hỏi thường chỉ được soạn thảo hạn chế trong phạm vi của từng giảng viên, sau đó áp dụng ngay cho các kỳ thi không qua một lần thử nghiệm nên thường hàm chứa những yếu tố chủ quan nhất định. Để có một bộ câu hỏi TNKQ tốt trong một đề thi, cần phải thử nghiệm để đánh giá tất cả các khía cạnh của câu hỏi nhằm đảm bảo phép đo (bộ đề câu hỏi) của chúng ta có tính giá trị và có thể tin cậy được.

Mục đích của nghiên cứu này là xây dựng bộ đề trắc nghiệm khách quan nhiều lựa chọn dùng đánh giá chất lượng học tập, bổ sung thêm vào hình thức thi viết truyền thống nhằm hoàn thiện quá trình dạy và học môn học Hoá Đại cương cho sinh viên khoa Hoá, trường Đại học Sư phạm thành phố Hồ Chí Minh.

## 2. Phương pháp nghiên cứu

### 2.1. Xây dựng bộ đề thi trắc nghiệm Hoá Đại cương

Trên cơ sở lí luận của Haladyna (1997) và Bloom (1956), chúng tôi đã xây dựng mỗi câu trắc nghiệm trên hai yếu tố đánh giá là *nội dung* và *hành vi nhận thức*.

Chúng tôi đã thực hiện 6 bước như sau:

#### **Bước 1:** Phân tích nội dung cần được kiểm tra

Căn cứ vào chương trình học, nội dung kiểm tra là tất cả những nội dung được học trong bốn chương, được phân thành bốn loại theo quan điểm của Haladyna (1997):

- ◊ Những sự kiện
- ◊ Những khái niệm
- ◊ Những luận điểm, qui tắc, qui luật, nguyên lí.
- ◊ Qui trình sản xuất trong hoá học.

#### **Bước 2:** Xác định mục tiêu môn học Hoá Đại cương A1

Căn cứ trên mục tiêu đào tạo bậc đại học và mục tiêu cụ thể của môn học Hoá đại cương, chúng tôi chọn các hành vi nhận thức cần được đánh giá theo thang đo kĩ năng nhận thức của Bloom gồm các mức độ từ thấp đến cao:

- ◊ Nhận biết
- ◊ Hiểu
- ◊ Áp dụng
- ◊ Phân tích, tổng hợp, đánh giá

**Bước 3:** Quy hoạch bài trắc nghiệm

Để bảo đảm tính giá trị của một bài trắc nghiệm, chúng tôi đã lập bảng quy hoạch chương trình, bảng quy hoạch theo bài, để đo lường được những điều cần đo đã xác định trong mục tiêu môn học.

**Bảng 1:** Quy hoạch bài trắc nghiệm cho chương trình Hoá Đại cương A1 (5 ĐVHT = 75 tiết)

MỤC TIÊU NỘI DUNG	1	2	3	4	Tổng số câu	Tỷ lệ %	%số tiết của chương /chương trình
	Nhận biết	Hiểu	Áp dụng	PT-TH-ĐG			
Chương I: Các khái niệm và định luật cơ bản trong hóa học.	10	11	10	3	34	18%	12%
Chương II: Cấu tạo nguyên tử	16	15	15	12	58	31%	24%
Chương III: Định luật tuần hoàn và Hệ thống tuần hoàn các nguyên tố hóa học.	4	6	9	9	28	15%	10%
Chương IV: Cấu tạo phân tử và liên kết hóa học.	12	17	18	18	65	35%	36%

**Bước 4:** Lựa chọn hình thức trắc nghiệm

Chúng tôi chọn hình thức trắc nghiệm nhiều lựa chọn (đa phương án hay multiple choice) với năm lựa chọn, vì những lí do sau:

- Có thể sử dụng để kiểm tra các kĩ năng nhận thức bậc cao, phù hợp với mục tiêu giáo dục đại học.
- Tránh được nhược điểm người học chỉ biết một phát biểu là sai nhưng có thể không biết phát biểu đúng là như thế nào (so với loại câu hỏi Đúng-sai).
- Với nhiều phương án lựa chọn, có thể đánh giá xu hướng người học thường sa vào những điểm yếu nào.
- Tỷ lệ đoán mò đúng của hình thức trắc nghiệm 5 lựa chọn là 20%, thấp nhất so với 3 lựa chọn (33,3%) và 4 lựa chọn (25%).

**Bước 5:** Xây dựng các “mồi nhử” cho từng câu

Để cho các mồi nhử được hấp dẫn và phù hợp với các lựa chọn, nghiên cứu đã căn cứ vào thực tế khách quan dạy và học môn học trong 5 năm (1999-2003). Các ý tưởng trong giáo trình môn học được sắp xếp đặt câu hỏi cho sinh viên giải quyết vấn

đề trong các giờ luyện tập. Tổng hợp các câu trả lời của sinh viên cho từng ý tưởng, lọc lựa câu trả lời đúng, số câu trả lời còn lại thường là những nhầm lẫn của sinh viên, chúng tôi dùng làm “mồi nhử”. “Mồi nhử” càng có tính thực tế càng dễ “nhử” bởi vì “cái sai ấy chính là cái sai có thực trong thực tế nhận thức của sinh viên” (Dương Thiệu Tống, 1995).

**Bước 6:** Phân bố các câu trắc nghiệm thành các đề thử nghiệm

Để bảo đảm tính khách quan trong các câu trả lời của sinh viên, nghiên cứu đã tìm cách hạn chế sự quay cốp, trao đổi đáp án của những sinh viên ngồi gần nhau bằng biện pháp phân bố các câu trắc nghiệm thành những đề tương đương có cùng chung số câu hỏi, nhưng vị trí các câu bị xáo trộn.

**2.2. Đối tượng tham gia nghiên cứu**

Đối tượng tham gia trong nghiên cứu này là 172 sinh viên gồm 101 sinh viên niên khoá 2003 – 2004 và 71 sinh viên niên khoá 2004 – 2005 thuộc hệ chính qui của hai lớp Hoá 1A và Hoá 1B, Khoa Hoá trường Đại học Sư phạm thành phố Hồ Chí Minh.

**2.3. Cách tiến hành nghiên cứu**

Đợt thử I gồm 200 câu trắc nghiệm chia thành hai đề tương đương A và B. Từ đề A xáo trộn các câu thành đề 1 và đề 4. Làm tương tự cho đề B.

Đợt thử II gồm 176 câu trắc nghiệm đã được chỉnh sửa ở đợt I. Khâu in ấn và photo để do chính tác giả nghiên cứu thực hiện.

**1. Công cụ nghiên cứu**

Công cụ nghiên cứu là bài trắc nghiệm gồm 200 câu hỏi trắc nghiệm năm lựa chọn đo lường các mức độ biết, hiểu, vận dụng, phân tích – tổng hợp – đánh giá các khái niệm, các nguyên lí, các qui luật cơ bản của hóa học.

Số liệu nghiên cứu căn cứ trên kết quả điểm của 101 bài (đợt I) và 71 bài trắc nghiệm (đợt II) thu được từ bảng trả lời theo thang 5 lựa chọn. Trước hết, nghiên cứu đã chấm các bài trắc nghiệm theo phương pháp đục lỗ thủ công để tiện kiểm soát sau này.

**2. Cách xử lí số liệu**

a) Việc chấm điểm từng bài, việc phân tích bài, phân tích câu trắc nghiệm được thực hiện trên máy vi tính, dùng phần mềm phân tích trắc nghiệm TEST.

b) Phân loại, sắp xếp các câu trắc nghiệm thành từng nhóm theo thứ tự độ khó và độ phân cách. So sánh số câu phân cách của hai bài trắc nghiệm đợt I và đợt II, dùng phần mềm SYSTAT 5.05.

3. Kết quả nghiên cứu

3.1. Kết quả chung của bài trắc nghiệm

Bảng 2: So sánh kết quả hai bài trắc nghiệm khảo sát đợt I và đợt II

Các giá trị thực nghiệm và giá trị lý thuyết	Đề A		Đề B	
	KS đợt I	KS đợt II	KS đợt I	KS đợt II
Số bài trắc nghiệm	49	35	52	36
Số câu	97	88	88	88
Số lựa chọn trong một câu	5	5	5	5
Trung bình toàn bài	66.490	63.914	53.500	54.694
Trung bình lý thuyết	59.400	52.800	53.400	52.800
Độ lệch chuẩn	9.881	8.276	8.173	10.099
Hệ số tin cậy	<b>0.848</b>	<b>0.825</b>	<b>0.786</b>	<b>0.865</b>
Độ khó bài TN	67.2%	72.6%	60.1%	62.2%
Độ khó vừa phải	60.0%	60.0%	60.0%	60.0%
Sai số tiêu chuẩn	3.858	3.458	3.778	3.712
Khoảng phân tán điểm số	47 → 85	47 → 80	37 → 69	34 → 74

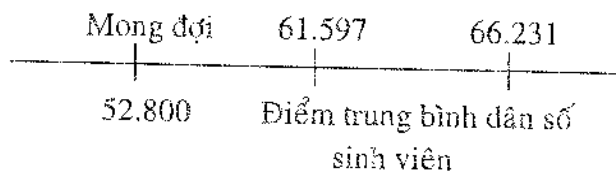
**Nhận xét:**

- Hệ số tin cậy bài trắc nghiệm của hai đề đều rất cao ( $\geq .85$ ) là rất tốt. Nếu so sánh với kết quả đợt I thì hệ số tin cậy của đề A đợt II tuy hơi bé hơn nhưng không đáng kể, trong khi ở đề B hệ số tin cậy bài trắc nghiệm đợt II lại lớn hơn đợt I. Như vậy, sau khi có chỉnh sửa, độ tin cậy của đề B đã tăng lên rõ rệt, trong khi đề A gần như không thay đổi.

- Trong 2 đợt khảo sát, độ khó đề A lần lượt là 67.2% và 72.6% đều  $> 60\%$  là độ khó vừa phải, trong khi ở đề B cả hai lần đều xấp xỉ 60%. Như vậy, bài trắc nghiệm của đề A tiếp tục dễ so với trình độ sinh viên, trong khi bài trắc nghiệm của đề B vẫn là vừa sức.

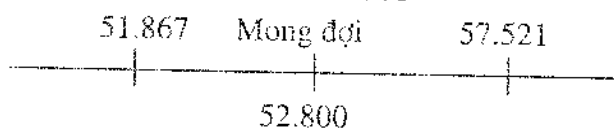
Để kiểm chứng tính chất trên, nghiên cứu dùng điểm trung bình của sinh viên. Khoảng ước lượng của trung bình dân số với độ tin cậy 95% như sau:

$$\begin{aligned} \text{Đề A: } & 63.914 - 1.96 \times 8.276 / \sqrt{49} = 63.914 - 2.317 = 61.597 \\ & 63.914 + 2.317 = 66.231 \end{aligned}$$



$$\text{Đề B: } 54.694 - 1.96 \times 10.099 / \sqrt{49} = 54.694 - 2.827 = 51.867$$

$$54.694 + 2.827 = 57.521$$



Điểm trung bình dân số sinh viên

Như vậy, trung bình lí thuyết của đề A nằm ngoài, về phía dưới khoảng ước lượng trung bình dân số sinh viên, bài làm sinh viên có điểm cao hơn mong đợi. Trong khi đó, trung bình lí thuyết của đề B nằm gần giữa khoảng ước lượng trung bình dân số sinh viên, bài làm sinh viên có điểm xấp xỉ điểm mong đợi. Vậy ở đợt II bài trắc nghiệm của đề A tiếp tục để so với trình độ sinh viên, trong khi bài trắc nghiệm của đề B vẫn là vừa sức.

### 3.2. Phân tích câu trắc nghiệm theo độ khó và độ phân cách

Kết quả kiểm nghiệm ý nghĩa của hệ số Rpbis cho thấy trong tổng số 176 câu của hai đề khảo sát đợt II có:

- + 35 câu phân cách tốt (xác suất ý nghĩa < .01);
- + 32 câu phân cách khá (xác suất ý nghĩa < .05);
- + 35 câu phân cách tạm chấp nhận (độ phân cách khá gần với mức có ý nghĩa < .05);
- + 56 câu phân cách kém và không phân cách;
- + 18 câu độ phân cách có giá trị âm.

Trong đó có:

\* 16 câu phân cách kém và phân cách âm sau khi chỉnh sửa trở thành phân cách dương có ý nghĩa mức < .01 và < .05. Một trong số các câu đó là:

Câu B58: (mã số B60 ở đợt thử I)

Câu số	Đợt thử I						Đợt thử II					
	B60						B58					
Lựa chọn	A	B*	C	D	E	Missing	A	B	C	D*	E	Missing
Tần số	4	3	7	26	7	5	1	0	3	20	10	2
Tỉ lệ %	8.5	6.4	14.9	55.3	14.9		2.9	0.0	8.8	58.8	29.4	
Rpbis	0.12	-0.32	-0.06	0.46	-0.27		-0.15	NA	-0.33	0.63	-0.31	
Mức ý nghĩa	NS	< .05	NS	< .01	NS		NS	NA	NS	< .01	NS	
Độ khó	Mean (câu) = 0.058						Mean (câu) = 0.556					

**Nhận xét:**

Số liệu cho thấy ở đợt I, đây là câu phân cách âm ( $Rpbis = -0.318$ ), đáp án B chỉ có 3 sinh viên chọn trong khi lựa chọn D có đến 26 sinh viên chọn với độ phân cách dương phản ảnh ngược. Missing là 5, tương đối lớn, có đến 5 sinh viên bỏ không làm câu này. Câu ở mức độ quá khó, mean (câu) = 0.058.

Xem lại nội dung câu:

*Hợp chất nào dưới đây có liên kết Hydro yếu nhất?*

- a)  $NH_3$       \*b)  $H_2O$       c)  $HF$       d)  $CH_4$       e) Tất cả đều đúng

Nghiên cứu đã phát hiện rằng câu hỏi thiếu chính xác, bởi vì trong 4 hợp chất đưa ra để lựa chọn, có  $CH_4$  là không cho liên kết hydro, còn 3 hợp chất kia  $NH_3$ ,  $H_2O$  và  $HF$  giữa các phân tử của chúng đều có liên kết hydro. Vì vậy việc so sánh "yếu nhất" là không rõ ràng và đáp án B chọn  $H_2O$  là hợp chất mà liên kết hydro giữa những phân tử của chúng là yếu nhất là thiếu chính xác. Nghiên cứu đã sửa lại câu hỏi, đồng thời đổi lại đáp án cho phù hợp như sau:

*Hợp chất nào dưới đây không có liên kết hydro giữa những phân tử của chúng?*

- a)  $NH_3$       b)  $H_2O$       c)  $HF$       \*d)  $CH_4$       e) Tất cả đều đúng

Ở đợt II, câu trắc nghiệm kèm những sửa chữa như trên đã đạt những kết quả khả quan. Câu trở thành phân cách dương có ý nghĩa mức  $< .01$ , câu rất tốt với độ phân cách tăng lên đáng kể  $Rpbis = 0.632$ . Ba lựa chọn A, C và E có  $Rpbis$  âm là khá tốt với tỷ lệ % chọn tương đối được, chỉ có lựa chọn B không có ai chọn. Missing là 2, chỉ có 2 sinh viên bỏ không làm câu này. Câu có độ khó vừa phải là 55.6%. Nhìn chung việc sửa chữa câu ở đợt II đã có tác dụng đáng kể.

Bên cạnh đó cũng có:

\* 26 câu có độ phân cách tốt và khá ở đợt thử I nay trở thành câu phân cách kém hoặc phân cách âm ở đợt thử II

\* Có 7 câu không phân cách với độ khó là 100% ở đợt thử II cũng là những câu quá dễ với độ phân cách kém ở đợt thử I, những câu này nghiên cứu quyết định loại bỏ.

\* 12 câu có kết quả phân cách đợt II tiến bộ, tăng lên hơn so với đợt I nhưng vẫn chưa đạt được chuẩn mong đợi.

\* 29 câu không có vấn đề gì sửa chữa ở đợt I, giữ nguyên áp dụng cho đợt II nhưng vẫn không đem lại hiệu quả tốt hơn.

#### 4. Bàn luận và kết luận

Quá trình phân tích số liệu ở hai đợt thử cho thấy cả hai đề đều đạt hệ số tin cậy khá tốt. Khi phân tích độ khó của từng câu, hầu như cả hai đề đều có đủ từ rất dễ đến rất khó, tuy nhiên số câu khó ít còn số câu dễ chiếm nhiều hơn, đến 55%. Các câu dễ thường thuộc mức độ biết, còn câu khó ở mức độ nhận thức cao hơn là hiểu, vận dụng, phân tích - tổng hợp - đánh giá.

Độ phân cách của các câu trắc nghiệm gồm đủ bốn mức độ từ tốt, khá đến tạm được và kém. Cả hai đề ở đợt II cho phép đánh giá tốt trình độ học tập của sinh viên, nhất là đề B.

Kết quả cũng thể hiện rõ nét rằng hệ số tin cậy của bài trắc nghiệm ở đợt II tăng lên đáng kể so với đợt I, từ 0.786 tăng lên 0.865 đối với đề B, trong khi đề A gần như không thay đổi bao nhiêu,  $R_{pbis}$  đợt II = 0.825 xấp xỉ với  $R_{pbis}$  đợt I = 0.848. Với trị số này, độ tin cậy của hai đề trắc nghiệm đợt II ở mức rất tốt (> .85). Điều này là hiển nhiên bởi vì tính ổn định của bài trắc nghiệm đợt II tốt hơn so với đợt I như đã bàn luận ở trên.

Điều đáng ghi nhận là qua hai đợt thử có 16 câu phân cách kém và phân cách âm ở đợt thử I trở thành câu phân cách khá và tốt ở đợt thử II, đã chứng tỏ rằng việc phân tích câu trắc nghiệm tìm lí do cần sửa chữa, xem xét nội dung và quyết định sửa chữa của nghiên cứu đã mang lại hiệu quả khá tốt.

Mặc khác, nghiên cứu đã phát hiện có 26 câu phân cách tốt và khá ở đợt I nay trở thành câu phân cách kém hoặc phân cách âm.

Ngoài ra, nghiên cứu cũng ghi nhận có 29 câu không có biến chuyển đáng kể qua hai đợt khảo sát. Đa số chúng là những câu thuộc loại khó, quá khó hoặc dễ, quá dễ. Sau khi phân tích các số liệu, đọc lại nội dung, nghiên cứu tìm thấy những nguyên nhân sau:

- Các "mồi nhử" và đáp án đúng ở dạng kiến thức tổng quát về hoá học, chứ không gắn trực tiếp với bài học. Vì thế mà sinh viên khó phân biệt hết ý nghĩa sâu xa bên trong và biểu hiện bên ngoài của một khái niệm, qui luật cơ bản trong hoá học, dẫn đến đoán mò kể cả sinh viên giỏi và sinh viên kém.

- Các "mồi nhử" quá "hấp dẫn", sinh viên khi đọc lướt qua sẽ khó phân biệt ra sự khác nhau rất nhỏ giữa đáp án đúng và lựa chọn sai.

- "Mồi nhử" soạn dở không thu hút được sinh viên, trong khi đáp án đúng lại bộc lộ quá rõ ý đúng làm cho câu có đến 2, 3 trên tổng số 4 lựa chọn không có ai chọn.

- Do người soạn câu trắc nghiệm có những sơ sót như:



- + Câu lựa chọn quá dài làm cho sinh viên tưởng đó là đáp án và chọn ngay mà không kịp đọc các lựa chọn sau.
- + Người soạn trắc nghiệm đã hành văn không tối làm sinh viên bị chỉ phối mà lựa chọn sai đáp án.
- + Kiểm tra không hết lỗi chính tả, công thức hóa học, lỗi in sai.

Từ 185 câu qua hai đợt thử nghiệm cứu đã chọn lọc được 166 câu. Các câu này đều rơi đủ vào các mức độ biết (62 câu), hiểu (37 câu), áp dụng (50 câu) và phân tích-tổng hợp - đánh giá (27 câu). Nếu xếp theo độ khó, trong 166 câu chọn lọc có 11 câu quá khó, 34 câu khó, 34 câu đạt độ khó vừa phải, 44 câu dễ và 49 câu quá dễ. Nhìn chung, bài trắc nghiệm chưa được cân đối ở các mức độ, số câu khó ít, phần lớn câu thuộc loại dễ và tập trung nhiều ở mức độ biết và áp dụng. Đây là điều hiển nhiên, vì câu trắc nghiệm càng khó, càng ở mức độ nhận thức cấp cao thì càng khó soạn thảo. 166 câu được chọn là những câu đạt tiêu chuẩn về mặt nội dung, sẽ đóng góp vào ngân hàng đề thi trắc nghiệm của môn học.

Tóm lại, việc soạn thảo các câu hỏi trắc nghiệm nhiều lựa chọn, thử nghiệm, chọn lọc và tiến đến thành lập ngân hàng lưu trữ đề thi các môn hóa học là công việc đòi hỏi nhiều thời gian và nỗ lực đầu tư.

Điều làm chúng ta e ngại nhất là đòi hỏi người thầy giáo phải mất nhiều công sức và thời gian để soạn thảo, thử nghiệm và phân tích các câu trắc nghiệm. Để có thể đạt các chỉ số thống kê về độ khó, độ phân cách theo chuẩn đánh giá, một câu trắc nghiệm thường phải được thử vài lần.

Tuy có những khó khăn như trên, trắc nghiệm thành quả học tập rất đáng được sử dụng ở các trường đại học Việt Nam nhất là đối với các bộ môn khoa học đại cương, khi số lượng sinh viên ngày càng đông với nhiều hệ đào tạo trong cùng thời điểm. Mặt khác, cùng với sự phát triển của kĩ thuật máy tính, nhiều chương trình xử lí thống kê được viết sẵn, công việc phân tích bài và câu trắc nghiệm; công việc chấm điểm và in ấn kết quả trở nên dễ dàng. Như vậy, trắc nghiệm thực sự là công cụ mạnh để đánh giá thành quả học tập của học sinh và cả quá trình giảng dạy của thầy giáo.

Muốn sử dụng được trắc nghiệm khách quan nhiều lựa chọn làm công cụ đánh giá chất lượng học tập, nhất thiết phải xây dựng ngân hàng đề thi trắc nghiệm. Về lâu dài, 166 câu trắc nghiệm cung cấp từ nghiên cứu này sẽ là vốn khởi đầu của ngân hàng lưu trữ đề thi trắc nghiệm môn học Hoá Đại cương.

### Ý kiến đề xuất

Để nâng cao chất lượng của quá trình thi cử, không thể cứ duy trì mãi lối ra đề thi, tổ chức thi và chấm bài theo lối thủ công truyền thống, cần phải tin học hoá quá trình thi cử. Công việc này chỉ có thể thực hiện được nếu được sự hỗ trợ của lãnh đạo nhà trường. Trên cơ sở đó, nghiên cứu xin đưa ra một số kiến nghị sau đây:

- 1) Hỗ trợ cho Khoa một phòng máy chuyên phục vụ cho qui trình thi và kiểm tra gồm một máy chủ và tối thiểu 40 máy con, để có thể kiểm tra 40 sinh viên trong một ca.
- 2) Tổ chức tập huấn, cung cấp kiến thức về khoa học đánh giá và đo lường thành quả học tập cho đội ngũ giảng viên các trường đại học.
- 3) Cần đưa thêm vào chương trình đào tạo các ngành sư phạm chuyên đề về “Trắc nghiệm và đo lường thành quả học tập” cũng như các “Phương pháp xử lý thống kê trong khoa học giáo dục”.

### TÀI LIỆU THAM KHẢO

- [1] Bloom, B. S, Madau, G. F., & Hastings, J. T. (1981), *Evaluation to improve learning*. New York: McGraw-Hill.
- [2] Dương Thiệu Tống (1995), *Trắc nghiệm và đo lường thành quả học tập (Phương pháp thực hành)*. Đại học Tổng hợp Tp. Hồ Chí Minh.
- [3] Gronlund, N. E. (1990), *Measurement and evaluation in teaching*. Macmillan Publishing Company.
- [4] Haladyna, T. M. (1994). *Writing test items to evaluate higher order thinking*. Needham Heights, MA: Allyn & Bacon.
- [5] Scoller, K. M. & Prosser, M. (1994), Students' experiences in studying for Multiple Choice Question Examination. *Studies in Higher Education*, 19 (3), tr. 267 – 279.

**Tóm tắt:**

**Đánh giá kết quả học tập môn học Hóa Đại cương  
bằng hình thức trắc nghiệm khách quan nhiều lựa chọn**

Nghiên cứu này nhằm giới thiệu cách xây dựng bộ đề thi gồm 200 câu hỏi trắc nghiệm khách quan nhiều lựa chọn, sử dụng làm công cụ đánh giá chất lượng học tập môn học Hóa Đại cương. Bộ đề trên được thử nghiệm hai lần với 172 sinh viên năm thứ nhất khoa Hóa, trường Đại học Sư phạm thành phố Hồ Chí Minh. Các chỉ số độ tin cậy, độ khó, độ phân cách của mỗi câu trắc nghiệm được thu thập và phân tích để đánh giá và chỉnh sửa khi cần thiết.

**Abstract:**

**Evaluating students' achievements in the course of General Chemistry  
through multiple choice question tests**

This study suggests a way to develop a test sample with 200 multiple-choice questions (MCQs) as a tool to evaluate students achievements in the General Chemistry course. This test sample was piloted two times (the first test and the second test) on 172 first-year students in the Chemistry Department of Ho Chi Minh City University of Education. The reliability, index of difficulty and discrimination of each MCQ were collected and analyzed in order to assess and improve the questions when necessary.