

BỘ GIÁO DỤC VÀ ĐÀO TẠO

ChuvanBien.vn
Chấp nhận tư tưởng lại

TÀI LIỆU TẬP HUẤN

**TĂNG CƯỜNG NĂNG LỰC THỰC HIỆN KIỂM TRA, ĐÁNH GIÁ
THEO CHƯƠNG TRÌNH GIÁO DỤC PHỔ THÔNG 2018
CẤP TRUNG HỌC PHỔ THÔNG**

MÔN VẬT LÝ

ChuvanBien.vn
Chấp nhận tư tưởng lại

Hà Nội, năm 2024

MỤC LỤC

Phần I. MỘT SỐ VẤN ĐỀ CHUNG VỀ ĐÁNH GIÁ NĂNG LỰC CỦA HỌC SINH

.....3

1. Khác biệt giữa quan niệm cũ và mới về đánh giá.....3

2. Một số nguyên tắc đánh giá.....3

3. Bài kiểm tra và bài thi4

LH Zalo 0985.82.9393 nh n tài li u, khoá h c h t l ng mi n phí



Phần I. MỘT SỐ VẤN ĐỀ CHUNG VỀ ĐÁNH GIÁ NĂNG LỰC CỦA HỌC SINH

Nghị quyết số 29-NQ/TW ngày 4/11/2013 Hội nghị Trung ương 8 khóa XI về đổi mới căn bản, toàn diện giáo dục và đào tạo nêu rõ định hướng về kiểm tra, đánh giá học sinh: “Việc thi, kiểm tra và đánh giá kết quả giáo dục, đào tạo cần từng bước theo các tiêu chí tiên tiến được xã hội và cộng đồng giáo dục thế giới tin cậy và công nhận. Phối hợp sử dụng kết quả đánh giá trong quá trình học với đánh giá cuối kỳ, cuối năm học; đánh giá của người dạy với tự đánh giá của người học; đánh giá của nhà trường với đánh giá của gia đình và của xã hội”.

Chương trình giáo dục phổ thông 2018 bảo đảm phát triển phẩm chất và năng lực người học thông qua nội dung giáo dục với những kiến thức, kỹ năng cơ bản, thiết thực, hiện đại; hài hoà đức, trí, thể, mỹ; chú trọng thực hành, vận dụng kiến thức, kỹ năng đã học để giải quyết vấn đề trong học tập và đời sống; tích hợp cao ở các lớp học dưới, phân hoá dần ở các lớp học trên; thông qua các phương pháp, hình thức tổ chức giáo dục phát huy tính chủ động và tiềm năng của mỗi học sinh, các phương pháp đánh giá phù hợp với mục tiêu giáo dục và phương pháp giáo dục để đạt được mục tiêu đó.

Mục tiêu đánh giá kết quả giáo dục là cung cấp thông tin chính xác, kịp thời, có giá trị về mức độ đáp ứng yêu cầu cần đạt của chương trình và sự tiến bộ của học sinh để hướng dẫn hoạt động học tập, điều chỉnh các hoạt động dạy học, quản lí và phát triển chương trình, bảo đảm sự tiến bộ của từng học sinh và nâng cao chất lượng giáo dục.

Phương thức đánh giá bảo đảm độ tin cậy, khách quan, phù hợp với từng lứa tuổi, từng cấp học, không gây áp lực lên học sinh, hạn chế tổn kém cho ngân sách nhà nước, gia đình học sinh và xã hội.

1. Khác biệt giữa quan niệm cũ và mới về đánh giá

Các khái niệm về đánh giá đã có những thay đổi, đặc biệt liên quan đến giáo dục dựa trên năng lực. Dưới đây là bảng so sánh sự khác biệt giữa quan niệm cũ và quan niệm mới về đánh giá.

<i>Quan niệm cũ</i>	<i>Quan niệm mới</i>
Đánh giá kết thúc một giai đoạn trong bối cảnh chính thức được sử dụng làm hình thức đánh giá chính (đôi khi là duy nhất)	Đánh giá quá trình và không chính thức
Kiểm tra cuối học kỳ hoặc cuối năm	Đánh giá liên tục là một phần không thể thiếu của quá trình dạy học
Tham khảo tiêu chuẩn, so sánh kết quả học tập của học sinh này với học sinh khác như một chỉ số xếp hạng cuối cùng và cho vị trí/lựa chọn	Tham khảo tiêu chí, so sánh kết quả học tập của học sinh so với các tiêu chí/tiêu chuẩn đã xác định trước để cung cấp phản hồi và cải thiện hiệu suất
Kiến thức và khả năng nhớ lại nội dung, bị căng thẳng nặng nề.	Áp lực đè nặng lên quá trình học

2. Một số nguyên tắc đánh giá

Việc đánh giá trong giáo dục thường phải bảo đảm một số nội dung dưới đây:

Độ giá trị: Một đánh giá có giá trị khi nó thực sự đánh giá được những gì đặt ra (đo lường).

Độ tin cậy: Việc đánh giá tạo ra kết quả giống nhau trong những trường hợp khác nhau và với những người đánh giá khác nhau.

Tính xác thực: Đánh giá có thể được chứng minh là có liên quan đến công việc của cá nhân học sinh.

Khả năng tiếp cận: Đánh giá có thể được thực hiện trong điều kiện nó được thực hiện thường xuyên, bao gồm nhiều nhiệm vụ trong phạm vi điều kiện cho phép.

Hiệu quả: Các phương pháp đánh giá được sử dụng tránh dài dòng và trùng lặp không cần thiết.

Sự đầy đủ của phản hồi: Các kết quả được ghi lại và có sẵn cho người dùng một cách nhanh chóng, đầy đủ chi tiết để có thể sử dụng một cách tích cực.

Hiệu quả chi phí: Để đáp ứng tất cả các yêu cầu nêu trên và ở mức độ có thể, các thủ tục đánh giá được áp dụng phải có hiệu quả về mặt chi phí.

3. Bài kiểm tra và bài thi

Hiệu quả học tập của cá nhân có thể được đo lường bằng cách sử dụng các công cụ khác nhau, thường được gọi là bài kiểm tra. Bài kiểm tra là một tập hợp các nội dung được thiết kế để đo lường hiệu quả học tập của một học sinh hoặc một nhóm học sinh.

a) Chức năng của các bài kiểm tra

Các bài kiểm tra cung cấp thông tin hỗ trợ cho quá trình ra quyết định. Tuy nhiên, các bài kiểm tra khác nhau về loại thông tin cần tìm kiếm, loại quyết định có thể hỗ trợ và kết quả kiểm tra có thể phù hợp cho một số quyết định này nhưng không phù hợp với các quyết định khác.

- Sử dụng để giám sát và đưa ra các quyết định quản lý đối với các nhóm học sinh (ví dụ: trường học, chương trình dạy học, chương trình môn học, cấp huyện, tỉnh, quốc gia)

Chẳng hạn: Đưa ra các quyết định về điều chỉnh chương trình giáo dục hay chương trình môn học. Đánh giá các chương trình thử nghiệm và đổi mới. Cung cấp thông tin chung về kết quả thực hiện của hệ thống giáo dục tổng thể. Báo cáo cho phụ huynh và ban giám hiệu nhà trường về kết quả thực hiện của từng trường hoặc từng địa phương.

- Sử dụng để giám sát quá trình học tập: Sự tiến bộ của mỗi học sinh; thông báo cho học sinh về quá trình học tập của bản thân; thông báo cho giáo viên về các quyết định giảng dạy.

Ví dụ: Chẩn đoán những mặt mạnh và yếu của học sinh. Giám sát việc làm chủ và thông hiểu nội dung học tập mới. Điều chỉnh chương trình giảng dạy theo tiến độ dựa vào kết quả bài kiểm tra. Giám sát tác động của một bài học hoặc một đơn vị học tập. Xây dựng kế hoạch bài dạy dựa trên mức độ hiểu bài hiện tại của học sinh.

- Sử dụng để phân bổ tài nguyên và cơ hội giáo dục giữa các cá nhân.

Ví dụ: Dự đoán hoặc phân loại: Sắp xếp học sinh vào học tại trường sao cho phù hợp nhất; Xếp lớp học tập theo chương trình cụ thể. Lựa chọn: Xét tuyển vào các trường

đại học. Chứng nhận: Chứng nhận hoàn thành khóa học, môn học hoặc chứng nhận tốt nghiệp.

b) Lĩnh vực thường được kiểm tra

Học sinh thường được kiểm tra nhiều nhất là những lĩnh vực liên quan đến:

Kiến thức: Mức độ kiến thức của người học trong lĩnh vực kỹ năng.

Kỹ năng: Khả năng của người học để thực hiện các nhiệm vụ trong lĩnh vực kỹ năng.

Thái độ: Cách tiếp cận của người học đối với công việc, nhiệm vụ của mình.

c) Đề kiểm tra

Trong các loại bài kiểm tra sử dụng để đánh giá học sinh thì bài kiểm tra viết bằng đề kiểm tra là một hình thức thường được sử dụng nhiều trong các nhà trường.

Để xây dựng một đề kiểm tra chuẩn hóa, người ta thường thực hiện theo các bước sau: Xác định mục đích kiểm tra; xây dựng bản đặc tả kỹ thuật của đề kiểm tra; xây dựng câu hỏi; tổ chức kiểm tra thử; tổ chức đánh giá câu hỏi; xây dựng đề kiểm tra; tổ chức đánh giá đề kiểm tra; xây dựng hướng dẫn sử dụng và minh chứng.

Trong các bước trên, để làm tốt bản đặc tả kỹ thuật của đề kiểm tra, người biên soạn đề kiểm tra cần trả lời một số câu hỏi như sau:

- + Nội dung đề kiểm tra là gì?
- + Sẽ sử dụng dạng đề kiểm tra nào?
- + Bài kiểm tra dành cho những học sinh nào?
- + Thời gian kiểm tra là bao lâu?
- + Tổ chức kiểm tra như thế nào?
- + Câu hỏi và bài tập sẽ được chấm điểm như thế nào?
- + Có bao nhiêu câu hỏi cần được xây dựng?
- + Đặc điểm tâm trắc của các câu hỏi và đề kiểm tra là gì?

Xây dựng ma trận đề kiểm tra

Ma trận đề kiểm tra là một bảng biểu thị cách thức để đo lường các năng lực. Theo cách này, kế hoạch kiểm tra cung cấp tổng quan về tất cả bài kiểm tra và phương pháp kiểm tra sẽ được sử dụng và ma trận đề kiểm tra hiển thị nội dung của từng đề riêng biệt.

Các ma trận đề kiểm tra dùng để đánh giá năng lực (sử dụng kiến thức, kỹ năng...) chỉ ra chủ đề của bài kiểm tra, mức độ thành thạo nào được mong đợi cho mỗi chủ đề và cách các chủ đề được phân phối như thế nào trong bài kiểm tra. Do đó, ma trận đề kiểm tra là một công cụ rất quan trọng trong quá trình xây dựng một bài kiểm tra. Nội dung học tập được thể hiện trong cấu trúc đề kiểm tra. Cấu trúc đề kiểm tra trình bày các kiến thức và kỹ năng có trong bài kiểm tra và qua đó thể hiện được các hành vi khi làm bài của học sinh.

Một số chức năng cơ bản của ma trận đề kiểm tra là: (1) Giúp quản lý bài kiểm tra. (2) Được sử dụng dưới dạng: Một công cụ hoặc tài liệu hướng dẫn để xây dựng bài kiểm tra; một phương tiện để điều khiển (vì có mối quan hệ trực tiếp giữa các khía cạnh cần được đánh giá, bài kiểm tra và cách xếp hạng hoặc chấm điểm); một dụng cụ để kết hợp các phương pháp một cách phù hợp; một dụng cụ để kết nối các bài kiểm tra và bài kiểm tra lại; một sự

LH Zalo 0985.82.9393 nh n tài li u, khoá h c ch t l ng mi n phí

minh chứng với bên ngoài. (3) Một phương tiện để nâng cao hiệu quả giảng dạy của giáo viên và việc học tập của học sinh.

Giá trị của ma trận đề kiểm tra: Ma trận đề kiểm tra mang tính chuẩn hóa. Nó là điểm khởi đầu rõ ràng cho cả người ra đề và học sinh làm bài. Ma trận đề kiểm tra cung cấp thông tin về tỉ trọng của từng nội dung và là căn cứ để xây dựng các đề kiểm tra tương đương nhau, có thể dùng làm cơ sở để so sánh các đề kiểm tra với nhau. Ma trận đề kiểm tra là công cụ nâng cao về độ giá trị của đề kiểm tra, đó là đề kiểm tra có đo được những gì định đo hay không?

Cách thiết lập ma trận đề kiểm tra phụ thuộc vào nhiều yếu tố. Thực tế, không có một mô hình lý tưởng duy nhất nào cho ma trận đề kiểm tra. Tuy nhiên, một ma trận đề kiểm tra thường bao gồm hai (hoặc nhiều) chiều, trong đó quan trọng nhất là nội dung chủ đề và cấp độ nhận thức. Để xây dựng được ma trận đề kiểm tra có chất lượng tốt, người xây dựng ma trận đề kiểm tra cần trả lời các câu hỏi sau: Các tiêu chí khác nhau trong bài kiểm tra được phân bổ trọng số như thế nào? Làm thế nào để có thể bao hàm đủ các mục tiêu và chủ đề? Làm thế nào để có thể bao hàm đủ các cấp độ nhận thức từ bậc thấp đến bậc cao? Làm thế nào để có thể bao hàm đủ các tình huống ứng dụng dự kiến? Làm thế nào để có thể bảo đảm rằng bài làm thực sự là của học sinh?

Trong bài kiểm tra viết thường chỉ có 2 loại câu hỏi được sử dụng, đó là câu hỏi dạng trắc nghiệm khách quan (trắc nghiệm) và câu hỏi tự luận. Đối với mỗi loại câu hỏi đều có những ưu điểm và những hạn chế nhất định.

Các kiểu câu hỏi trắc nghiệm khách quan

(i) Câu ghép đôi (matching items) đòi hỏi thí sinh phải ghép đúng từng cặp nhóm từ ở hai cột với nhau sao cho phù hợp về ý nghĩa. Đối với loại câu hỏi này, người ra đề thường cho số yếu tố ở cột bên trái không bằng số yếu tố ở cột bên phải, để tránh thí sinh dùng phương pháp loại trừ.

(ii) Câu điền khuyết (fill-in-the-blanks): nêu một mệnh đề có khuyết một bộ phận, thí sinh phải nghĩ ra nội dung thích hợp để điền vào chỗ trống.

(iii) Câu trả lời ngắn (short answer): là câu trắc nghiệm đòi hỏi trả lời bằng nội dung rất ngắn.

(4i) Câu đúng -sai (true/false): đưa ra một nhận định, thí sinh phải lựa chọn một trong hai phương án trả lời để khẳng định nhận định đó là đúng hay sai.

(5i) Câu nhiều lựa chọn (multiple choice questions) đưa ra một nhận định và thường có 3 – 7 phương án trả lời, thí sinh phải chọn để đánh dấu vào một phương án đúng hoặc phương án tốt nhất.

Trong các kiểu câu trắc nghiệm đã nêu, kiểu câu đúng – sai và kiểu câu nhiều lựa chọn có cách trả lời đơn giản nhất. Câu đúng – sai cũng chỉ là trường hợp riêng của câu nhiều lựa chọn với 2 phương án trả lời. Đối với câu nhiều lựa chọn, nếu có n phương án trả lời thì xác suất thí sinh làm đúng là $1/n$. Trong các kiểu câu trắc nghiệm, kiểu câu nhiều lựa chọn được sử dụng phổ biến hơn cả vì chúng có cấu trúc đơn giản, dễ xây dựng thành bài thi, dễ chấm điểm. Trong khi soạn thảo 5 câu trắc nghiệm, người ra đề thường đưa vào các phương án “gây nhiễu” (distractor) và thường cố gắng làm cho các phương án gây nhiễu đều có vẻ “có lý” và “hấp dẫn”. Về nguyên tắc, đối với người có kinh nghiệm viết trắc nghiệm, một nội dung bất kỳ nào cần kiểm tra đều có thể được thể hiện vào một câu trắc nghiệm theo một kiểu nào đó.

Vì thế đối với tất cả các môn học người ta đều có thể viết câu hỏi trắc nghiệm. Tuy nhiên, do đặc thù của từng môn học mà việc viết trắc nghiệm cho môn này có thể khó hơn cho môn kia. Cần lưu ý rằng không phải bất cứ ai có kiến thức chuyên môn cũng viết được câu trắc nghiệm có chất lượng cao cho chuyên môn đó. Muốn viết câu hỏi trắc nghiệm tốt phải suy nghĩ sâu sắc về chuyên môn và tích lũy kinh nghiệm sau một thời gian thử nghiệm lâu dài. Một số người không có khả năng viết được câu trắc nghiệm tốt hoặc không hiểu sâu về các câu trắc nghiệm thường vội kết luận rằng trắc nghiệm chỉ đánh giá được khả năng nhớ tầm thường.

Ưu điểm và nhược điểm của câu hỏi trắc nghiệm

Tính hiệu quả của loại câu hỏi trắc nghiệm: Học sinh không lãng phí thời gian để đưa ra câu trả lời, giáo viên tiết kiệm thời gian chấm bài, phù hợp cho việc chấm điểm tự động. Nhưng loại câu hỏi này chỉ đánh giá sự tái tạo kiến thức và nếu xem xét kỹ hơn thì “đáp án đúng” đôi khi cũng không chính xác, thậm chí là còn sai và một số phương án nhiễu có thể không phải là không đúng. Việc xây dựng câu hỏi trắc nghiệm là một công việc khó khăn và nhiều bài kiểm tra trắc nghiệm không đáp ứng được các tiêu chí. Tuy nhiên, nếu câu hỏi trắc nghiệm được xây dựng phù hợp sẽ có nhiều ưu điểm và có thể được sử dụng để đánh giá tốt hơn so với mọi người vẫn nghĩ.

Ưu điểm lớn của câu hỏi trắc nghiệm là nó giúp tăng cường độ tin cậy của bài kiểm tra. Việc sử dụng loại câu hỏi này cho phép có nhiều câu hỏi hơn trong một bài kiểm tra. Ngoài ra, việc chấm điểm câu hỏi trắc nghiệm rất đáng tin cậy khi đã xác định được đáp án đúng. Hầu hết các nhược điểm của câu hỏi trắc nghiệm xuất hiện trong giai đoạn biên soạn câu hỏi, mặc dù một số nhược điểm không được bộc lộ cho đến khi chấm điểm bài kiểm tra.

Một lợi thế của bài kiểm tra trắc nghiệm là nó cho phép học sinh trả lời được nhiều câu hỏi trong một thời gian tương đối ngắn, do đó có thể bao phủ đáng kể nội dung học tập. Trung bình, các câu hỏi Đúng-Sai cần khoảng 50 giây để trả lời và một câu hỏi trắc nghiệm với 4-5 lựa chọn mất khoảng 75 giây¹. Một nhược điểm là một nội dung học tập có thể ít phù hợp với câu hỏi trắc nghiệm và điều này có thể phủ nhận lợi thế đã nói ở trên về độ bao phủ nội dung học tập. Tuy nhiên, câu hỏi trắc nghiệm phù hợp với nhiều mục đích hơn, cho dù chúng thường được sử dụng cho mục đích đánh giá việc tái tạo (nhớ) kiến thức trong thực tế.

Ưu điểm và nhược điểm của các câu hỏi tự luận

Câu hỏi tự luận là câu hỏi buộc thí sinh phải trả lời theo dạng mở, thí sinh phải tự mình trình bày ý kiến trong một bài viết để giải quyết vấn đề mà câu hỏi nêu ra.

Câu hỏi tự luận rất phù hợp với việc đánh giá lập luận, kiểm tra các kỹ năng nhận thức bậc cao (ví dụ: ứng dụng, phân tích, tổng hợp và đánh giá) và đối với các lĩnh vực chủ đề rộng và không có giới hạn rõ ràng, vì loại câu hỏi này ít hạn chế sự tự do của học sinh. Câu hỏi tự luận ít phù hợp hơn cho việc đánh giá kiến thức thực tế so với câu hỏi trắc nghiệm vì câu hỏi trắc nghiệm đem lại hiệu quả và độ tin cậy cao hơn. Câu hỏi tự luận cũng thích hợp hơn câu hỏi trắc nghiệm trong các bài kiểm tra đánh giá sự thành thạo trong sử dụng kỹ năng ngôn ngữ.

Câu hỏi tự luận đặt ra yêu cầu cao hơn về các kỹ năng ngôn ngữ của học sinh. Không phải lúc nào cũng có thể rõ liệu câu trả lời không chính xác là do không đủ kiến thức, không

¹ CITo

đủ thành thạo nội dung môn học hay do không đủ kỹ năng ngôn ngữ. Cũng rất khó đặt ra câu hỏi để mọi học sinh đều hiểu giống nhau. Các câu hỏi được biên soạn không rõ ràng có thể gây khó khăn cho việc chấm điểm, bởi vì có khả năng học sinh sẽ trả lời theo cách ngoài dự kiến. Kể cả những câu hỏi được biên soạn rõ ràng vẫn có thể gặp khó khăn khi chấm điểm. Người chấm điểm có thể sẽ không thống nhất với nhau về số điểm được chấm cho từng câu trả lời cụ thể. Cần nhiều thời gian để trả lời câu hỏi tự luận hơn so với câu hỏi trắc nghiệm. Do đó, trong cùng một khoảng thời gian làm bài, bài kiểm tra tự luận có ít câu hỏi hơn. Đồng thời, việc chấm điểm câu tự luận cũng mất nhiều thời gian hơn.

d) Đề thi

Độ giá trị của đề thi là thuật ngữ được dùng khi một đề thi đo lường chính xác những gì mà đề thi dự định đo. Trong lĩnh vực “tâm trắc” (trắc nghiệm tâm lý) và khảo thí giáo dục, độ giá trị nói đến mức độ đáp ứng của các cơ sở thực tiễn và các cơ sở lý thuyết với giả thuyết về điểm thô của đề thi kèm theo mục đích sử dụng của đề thi.

Độ giá trị của đề thi giúp đảm bảo rằng đề thi tuân theo các tiêu chuẩn chuyên môn và kỹ thuật nhất định để đo lường những kiến thức, kỹ năng, năng lực đề thi được thiết kế nhằm mục đích đo. Nói cách khác, độ giá trị cho chúng ta biết đề thi đo lường được bao nhiêu lượng kiến thức chúng ta dự kiến đo lường. Các phương pháp chính để biểu diễn độ giá trị của đề thi bao gồm độ giá trị tiêu chí, độ giá trị nội dung, độ giá trị bề mặt và độ giá trị cấu trúc (Brennan, 2006).

Để đảm bảo độ giá trị của đề thi, các câu hỏi cũng như đề thi cần phải đánh giá được năng lực, kiến thức dự kiến, đồng thời các câu hỏi thi phải đáp ứng được yêu cầu về độ tin cậy, độ phân biệt, độ khó.

Độ tin cậy của bài thi phản ánh mức độ nhất quán trong kết quả của các lần thi. Thường thì 3 đặc tính sau quyết định độ tin cậy của bài thi:

- + Sự tương quan lẫn nhau giữa các câu hỏi: mức độ tương quan càng lớn thì mối quan hệ giữa các câu hỏi càng mạnh do đó độ tin cậy của bài thi càng cao; đặc tính có liên quan tới độ phân biệt của câu hỏi thi;
- + Độ dài của bài thi - bài thi có nhiều câu hỏi thường có độ tin cậy cao hơn;
- + Nội dung của bài thi - nhìn chung nếu bài thi hướng đến đánh giá nhiều nội dung thì độ tin cậy thấp hơn bài thi hướng đến đánh giá 1 nội dung.

Về mặt lý thuyết, hệ số tin cậy nằm trong khoảng từ 0 (không tin cậy) đến 1 (tin cậy hoàn toàn). Tuy nhiên trong thực tế, khoảng tin cậy thường từ 0.5 đến 0.9.

Độ tin cậy	Ý nghĩa
Từ 0.90 trở lên	Độ tin cậy hoàn hảo; đây là mức độ tin cậy của các đề thi chuẩn hóa tốt nhất
0.80 - 0.90	Độ tin cậy rất tốt cho đề thi

LH Zalo 0985.82.9393 nh n tài li u, khoá h c ch t l ng mi n phí

0.70 - 0.80	-	Độ tin cậy tốt cho đề thi. Tuy nhiên có thể có một số câu hỏi cần phải được cải tiến.
0.60 - 0.70	-	Độ tin cậy tương đối thấp. Bài thi này cần phải có các phương pháp đánh giá khác hỗ trợ để đưa ra kết quả. Rõ ràng là một số câu hỏi trong đề thi này cần phải được thẩm định lại nhằm nâng cao độ tin cậy.
0.50 - 0.60	-	Khuyến cáo nên thẩm định lại đề thi, trừ khi đề thi quá ngắn (chỉ bao gồm từ 10 câu hỏi trở xuống). Nếu phải sử dụng kết quả bài thi thì nên sử dụng các phương pháp đánh giá khác bổ sung để đưa ra kết quả/nhận định chính xác.
Từ 0.50 trở xuống	-	Độ tin cậy là không đáp ứng yêu cầu. Kết quả bài thi không nên sử dụng để ra quyết định cuối cùng đồng thời cần phải thẩm định lại đề thi.

Độ khó của câu trắc nghiệm là tỉ lệ thí sinh trả lời đúng so với tổng số thí sinh tham gia trả lời câu hỏi đó, được sử dụng rộng rãi đối với các câu hỏi đúng-sai, nhiều lựa chọn. Theo Osterlind (1989), giá trị độ khó càng lớn thì câu hỏi càng dễ và độ khó của câu hỏi nên nằm trong khoảng từ 0,4 đến 0,8. Cũng cần phải xem xét câu trắc nghiệm đó khó hay dễ với đối tượng nào. Thông dụng nhất là tính độ khó của một câu trắc nghiệm dựa trên tỉ lệ phần trăm số người trả lời đúng câu trắc nghiệm đó so với tổng số người tham gia trả lời trắc nghiệm.

Độ khó của câu trắc nghiệm = (Số người trả lời đúng câu hỏi)/(Tổng số người tham gia trả lời)

Để xem xét một câu trắc nghiệm có độ khó bao nhiêu là phù hợp, phải tính xác suất làm đúng câu trắc nghiệm, xác suất này thay đổi tùy theo n, số lượng phương án lựa chọn trong mỗi câu trắc nghiệm, hay còn gọi là tỉ lệ may rủi (T).

$$T = (1/n).100\%$$

Theo Lâm Quang Thiệp (2008), về mặt lý thuyết (ĐKtb) của một câu hỏi trắc nghiệm có n phương án lựa chọn được xác định như sau: ĐKtb = (100% + T)/2

+ ĐK ≈ ĐKtb: câu trắc nghiệm được cho là tốt;

+ ĐK < ĐKtb: câu trắc nghiệm được cho là khó;

+ ĐK > ĐKtb: câu trắc nghiệm được cho là dễ.

Theo tác giả Lord (1952), độ khó lý tưởng của câu hỏi nhiều lựa chọn đáp ứng yêu cầu về độ phân biệt là:

Dạng thức	Độ khó lý tưởng
-----------	-----------------

Câu hỏi 1 phương án đúng/5 lựa chọn	0.70
Câu hỏi 1 phương án đúng/4 lựa chọn	0.74
Câu hỏi 1 phương án đúng/3 lựa chọn	0.77
Câu hỏi Đúng – Sai (1 phương án đúng/2 lựa chọn)	0.85

Khi thiết kế đề trắc nghiệm, một trong những yêu cầu của nó là có thể phân loại được những nhóm thí sinh có năng lực khác nhau như: khá, giỏi, trung bình.... Khả năng của câu trắc nghiệm có thể thực hiện được sự phân biệt đó được gọi là độ phân biệt. Một câu trắc nghiệm bảo đảm độ phân biệt là một câu trắc nghiệm tốt. Để thực hiện sự phân tích này, người ta so sánh kết quả làm bài của 2 nhóm thí sinh. Nhóm thí sinh có kết quả cao (27% số thí sinh làm bài test có tổng điểm bài test cao nhất) và nhóm khác có kết quả thấp (27% số thí sinh làm bài test có tổng điểm bài test thấp nhất). Sự chênh lệch giữa 2 nhóm trong việc lựa chọn các phương án trả lời được xem là hợp lý khi giá trị chênh lệch lớn hơn 0 đối với đáp án đúng. Ngược lại, các giá trị chênh lệch phải nhỏ hơn 0 đối với các phương án nhiễu còn lại. Các câu hỏi không đáp ứng tốt các yêu cầu trên thì cần được xem xét lại.

Độ phân biệt là khả năng mà câu trắc nghiệm phân loại được người học thành những nhóm có trình độ khác nhau. Sự phân biệt này mô tả chi tiết số người trả lời đúng nằm ở nhóm người đạt điểm cao ở toàn bài so với số người trả lời đúng nằm ở nhóm người đạt điểm thấp toàn bài. Theo Ebel (1965), các câu hỏi của bài kiểm tra nên có độ phân biệt bằng 0,3 hoặc cao hơn.

Cách tính độ phân biệt của câu hỏi trắc nghiệm: $D_{pb} = (N_c - N_t)/N$ trong đó:

- + N_c là số thí sinh trả lời đúng ở nhóm cao.
- + N_t là số thí sinh trả lời đúng ở nhóm thấp.
- + N là số thí sinh ở nhóm cao điểm hoặc thấp điểm ở bài trắc nghiệm chiếm khoảng 27% tổng số thí sinh.

Thang đánh giá độ phân biệt D_{pb} thực hiện như sau:

- + $D_{pb} \geq 0,4$: câu trắc nghiệm có độ phân biệt rất tốt.
- + $0,30 \leq D_{pb} \leq 0,39$: câu trắc nghiệm có độ phân biệt khá tốt.
- + $0,20 \leq D_{pb} \leq 0,29$: câu trắc nghiệm có độ phân biệt trung bình (chấp nhận được).
- + $D_{pb} \leq 0,19$: câu trắc nghiệm có độ phân biệt kém.