

VẬN DỤNG FASTTEST ĐỂ XÂY DỰNG NGÂN HÀNG CÂU HỎI ĐÁNH GIÁ NĂNG LỰC GIẢI QUYẾT VẤN ĐỀ CỦA HỌC SINH

NGUYỄN THỊ LAN PHƯƠNG*

Ngày nhận bài: 06/12/2016; ngày sửa chữa: 15/12/2016; ngày duyệt đăng: 19/12/2016.

Abstract: Based on Item response theory, the article gives instructions to apply FastTEST to develop the item bank for assessing problem solving competence of students. The process described in the article includes 6 steps (Designing Item bank software, compiling items, testing, collecting data tests, calibrating items and selecting items into Item bank). The FastTEST is proper for developing the item bank for assessing problem solving competence of students at primary schools and secondary schools.

Keywords: Item bank, learner's competence assessment, Item response Theory, calibrating item, estimating an examinee's ability; FastTEST.

Trong một bài viết khác về “Phát triển ngân hàng câu hỏi (CH) để đánh giá năng lực (NL) học sinh (HS) phổ thông” chúng tôi đã: Tổng quan lí thuyết ứng đáp câu hỏi (IRT) về các thuật toán ước tính các tham số của CH và NL của HS; Xác định vai trò quan trọng của ngân hàng câu hỏi (NHCH) trong quá trình giảng dạy và đánh giá; Đề xuất ý tưởng vận dụng IRT để xây dựng một NHCH nhằm đánh giá NL của HS phổ thông. Bài viết này mô tả quá trình hiện thực hóa ý tưởng đó bằng cách: vận dụng phần mềm FastTEST của ASC (Assessment Systems Corporation) [1] để xây dựng NHCH đánh giá NL giải quyết vấn đề (GQVĐ) cho HS.

1. Quy trình phát triển NHCH dựa vào IRT

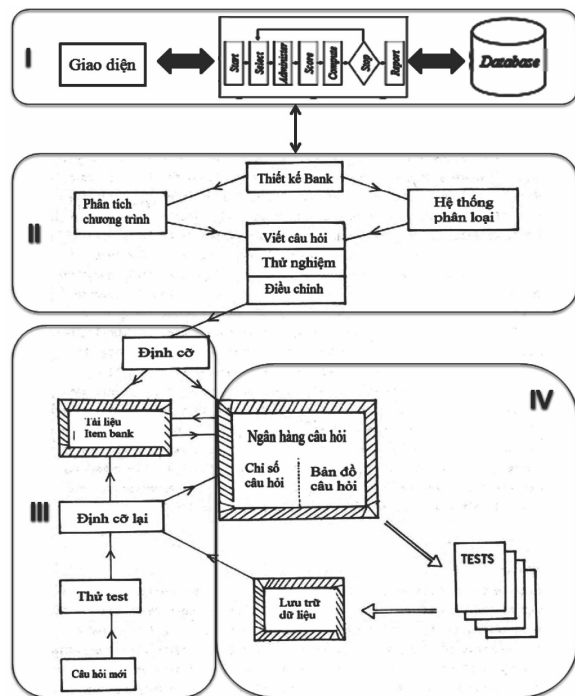
Theo David J. Weiss [2], phát triển một NHCH thường bao gồm 6 hợp phần chính là: (i) Xây dựng phần mềm NHCH; (ii) Biên soạn hệ thống CH test; (iii) Thiết kế test thử nghiệm; (iv) Thu thập dữ liệu test; (v) Định cỡ CH; (vi) Lựa chọn CH đưa vào NHCH. Trong đó, bốn hợp phần cuối có thể thực hiện trên cơ sở IRT và mô hình Rasch (xác suất P trả lời đúng CH thứ j có độ khó b_j của một người có NL θ được tính theo

công thức $P_j(\theta) = \frac{1}{1+e^{-(\theta-b_j)}}$, với e là cơ số của logarit tự nhiên [3].

Lê Xuân Tài và Đặng Hoài Phương [4], [5] đã đưa ra mô hình Trắc nghiệm thích nghi (TNTN) dựa trên cơ sở IRT. Theo đó, với thông số đầu vào là ước tính ban đầu về mức độ NL của thí sinh, có thể xem NHCH như là một thuật toán lặp gồm bốn bước: (i) Tất cả các CH chưa được lựa chọn sẽ là CH ứng cử cho lựa chọn tiếp theo dựa vào mức NL hiện tại của thí sinh; (ii) CH

tiếp theo phù hợp với NL hiện tại của thí sinh được đưa ra và thí sinh trả lời CH đó; (iii) Một ước lượng mới về năng lực được tính toán; (iv) Quay lại bước (i) nếu điều kiện ngừng chưa thỏa mãn (sai số ước tính chưa đáp ứng yêu cầu). Trong đó, thuật toán lựa chọn CH tiếp theo phù hợp với khả năng hiện tại của thí sinh ở bước (ii) là quan trọng nhất.

Từ đó, chúng tôi đề xuất quy trình phát triển NHCH theo 4 giai đoạn (hình 1):



Hình 1. Quy trình 4 giai đoạn phát triển NHCH

* Viện Khoa học Giáo dục Việt Nam

- *Giai đoạn I. Xây dựng phần mềm NHCH:* Gồm hợp phần Xây dựng phần mềm NHCH, với các hoạt động chủ yếu là: Thiết kế hệ thống quản lý CH, quản lý thí sinh và quản lý NHCH; Lập chương trình (theo thuật toán TNTN của Lê Xuân Tài và Đặng Hoài Phương); và Chạy thử chương trình.

- *Giai đoạn II. Thiết kế và thử nghiệm CH:* Gồm ba hợp phần: Biên soạn CH test, Thiết kế test thử nghiệm và Thu thập dữ liệu CH test, với các hoạt động chủ yếu là: Phân tích chương trình giáo dục; Thiết kế ma trận test; Biên soạn câu hỏi; Thiết kế test; Thiết kế mẫu thử nghiệm; Thu thập dữ liệu đáp ứng cho các CH.

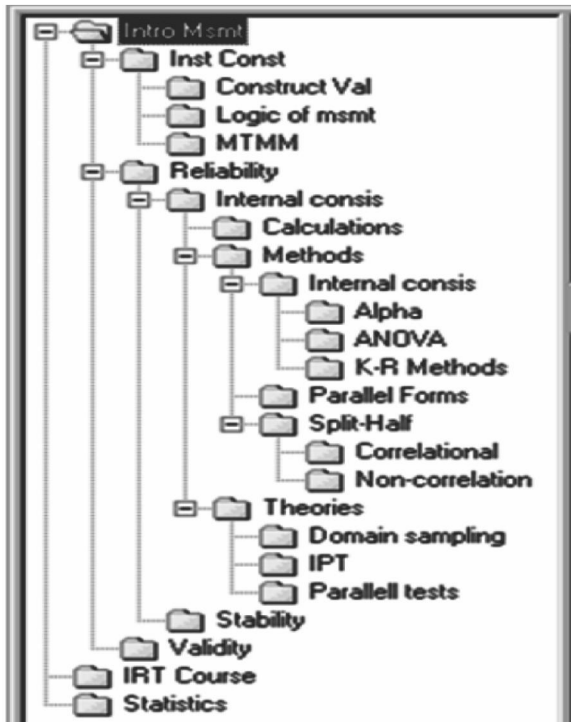
- *Giai đoạn III. Định cỡ CH:* Gồm hợp phần Định cỡ CH test, trong đó có hai hoạt động là: Định cỡ các CH gốc; và Định cỡ các CH mới bổ sung.

- *Giai đoạn 4. Đưa CH vào/ra NHCH:* Gồm hợp phần Lựa chọn CH cho NHCH, với các hoạt động chủ yếu là: Nhập vào NHCH các CH đáp ứng yêu cầu; Lưu trữ, cập nhật và hiệu chỉnh chúng; Xuất ra các bài test đánh giá NL thí sinh; Đưa các CH mới về cùng thang đo; Đưa CH mới đáp ứng yêu cầu vào NHCH.

2. Tóm lược quá trình phát triển NHCH đánh giá năng lực Giải quyết vấn đề của học sinh phổ thông

2.1. Phần mềm FastTEST

FastTEST 2.0 [1] là một hệ thống 32-bit, có chức năng tập hợp, lắp ráp các CH để tạo các bài test. Nó



Hình 2. Cấu trúc phân cấp của FastTEST

cho phép tạo NHCH có cấu trúc thư mục theo chủ định của người sử dụng, tạo ra các bài test theo các đặc tính IRT hoặc theo thống kê cụ thể của người sử dụng. FastTEST có nhiều cách lựa chọn CH test đáp ứng cao nhu cầu người sử dụng.

FastTEST có 6 cấp độ chức năng (xem hình 2). Cấp độ đầu tiên là ba ngân hàng *Intro Msmt*, *IRT Course*, và *Statistics*. Ngân hàng *Intro Msmt* có 5 cấp nhánh, còn hai ngân hàng kia thì không. Cấp độ tiếp theo của *Intro Msmt* là *Inst Const*, *Realibility* và *Stability*. Riêng *Realibility* được phân cấp sâu tới 4 mức, phản ánh cấu trúc phức tạp của các thuật toán ước tính độ tin cậy của test thích nghi. Sau đây mô tả ba chức năng chính của FastTEST.

a) *Thiết lập không gian làm việc (Workspaces).* Đây là nơi lưu giữ CH, bài test của NHCH. Người sử dụng có thể tạo nhiều *Không gian làm việc* khác nhau, dễ dàng di chuyển giữa các CH, bài test trong một không gian làm việc. Các chức năng cơ bản trong *Workspaces* gồm: Mở, tạo một không gian làm việc; Thiết lập CH; Xuất, nhập NHCH; Di chuyển bài test;...

b) *Tạo và làm việc với các NHCH (Item banks).* Sau khi tạo ra một không gian làm việc, tiếp theo là mở một hoặc nhiều ngân hàng. Có thể truy cập, tạo ra các ngân hàng thông qua thanh điều khiển hoặc Menu. Các chức năng cơ bản của *Item banks* gồm: Tạo cấu trúc Ngân hàng (ma trận test); Thêm CH vào ngân

@@ID=Sinh học\lop11\bai1

@@BT

Câu hỏi viết ở đây

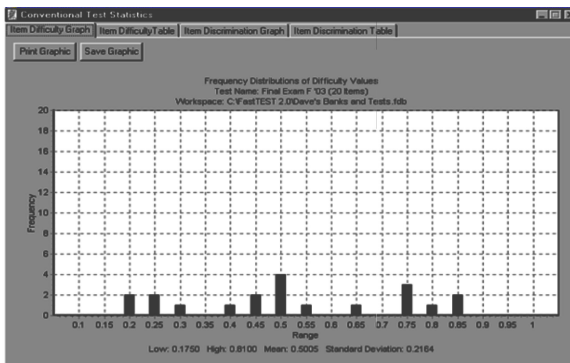
@@ET

Hình 3

hàng theo bốn cách (tạo trực tiếp bằng *Item Editor*; Nhập nội dung bằng *copy* và *paste*; Nhập từng CH như một tập tin RTF; Nhập nhiều CH từ file Word/Text) bằng *Item Text Import Program*, thông qua các đoạn mã để FastTEST nhận dạng nơi bắt đầu - kết thúc của CH (xem hình 3); Chỉnh sửa, cập nhật CH;...

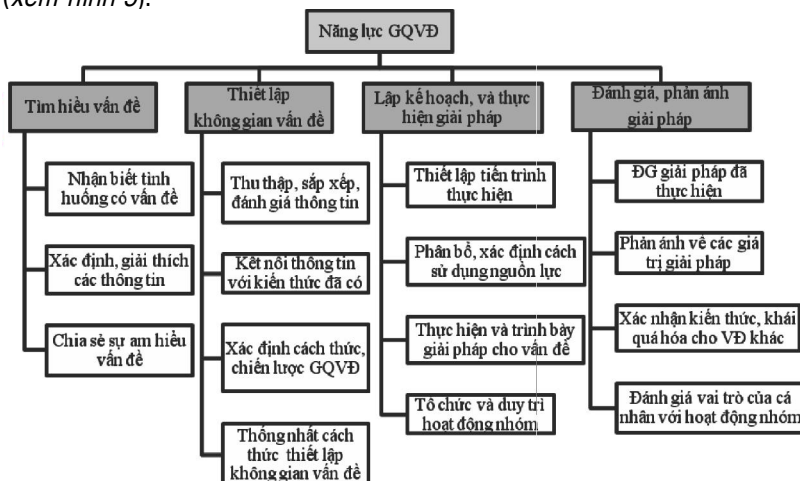
c) *Lắp ráp bài test.* Dựa vào các CH có sẵn, FastTEST cho phép lựa chọn CH cho bài test theo bốn cách: Duyệt các CH từ cấu trúc NHCH; Lựa chọn CH ngẫu nhiên; Lựa chọn CH bằng công cụ *Search for item*; Chọn các phần của ngân hàng.

FastTEST cho phép: Chẩn đoán các đặc điểm thống kê của bài test thông qua thống kê tham số CH và bản đồ đặc trưng CH (xem hình 4); Sắp xếp lại các CH (sắp xếp thủ công hoặc ngẫu nhiên); Tạo form trộn để thành nhiều mã.



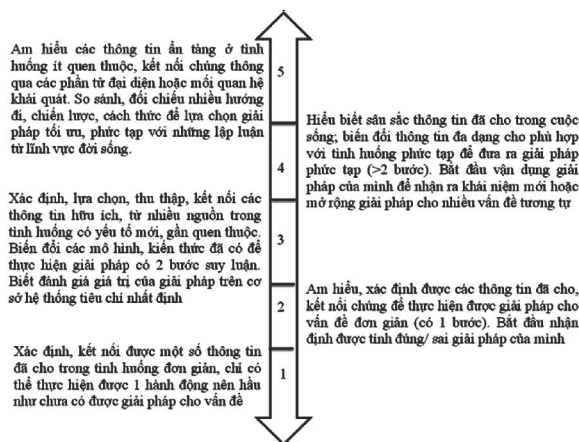
Hình 4. Bản đồ chẩn đoán độ khó các CH test

2.2. Thiết kế và thử nghiệm CH. Theo Nguyễn Thị Lan Phương [6], NL GQVĐ được cấu trúc gồm 4 kĩ năng (Tìm hiểu vấn đề, Thiết lập không gian vấn đề, Lập kế hoạch và thực hiện giải pháp, Đánh giá và phản ánh giải pháp) và thể hiện cụ thể qua 15 hành vi (xem hình 5).



Hình 5. Cấu trúc NL GQVĐ

NL GQVĐ được phát triển cho HS thông qua nội dung chương trình môn Toán, theo 5 mức độ được mô tả chi tiết ở hình 6.



Hình 6. Đường phát triển NL GQVĐ thông qua môn Toán

Mỗi CH sẽ đo lường từng hành vi, ứng với từng thành tố của NL GQVĐ (hình 5), độ khó của CH dựa vào các mức độ của đường phát triển NL (hình 6). Nó sẽ được xoay quanh ba trục: Kiểu tình huống thực tiễn (cuộc sống cá nhân, nhà trường, cộng đồng, khoa học); Nội dung toán (Số học và đại số, Hình học và đo đại lượng; Thống kê và xác suất); Hành vi NL GQVĐ cùng mức độ phát triển NL.

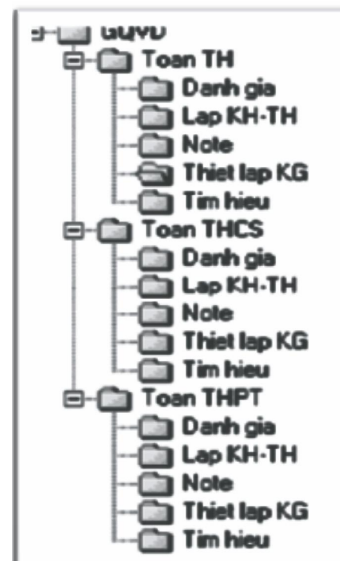
Ví dụ: *Giá điện trong mỗi tháng cho hộ gia đình được tính theo chỉ số điện tiêu thụ (chưa thuế) như bảng giá hình 7.*

100 chỉ số đầu:	1242 đồng
50 chỉ số tiếp theo:	1369 đồng
50 chỉ số tiếp theo:	1734 đồng
50 chỉ số tiếp theo:	1877 đồng

Hình 7

Nếu một gia đình tháng này dùng hết 120 chỉ số điện thì tổng số tiền phải thanh toán là bao nhiêu, với 10% thuế giá trị gia tăng.

CH này được thiết kế theo: Tình huống “Cộng đồng”; Nội dung “Số học và đại số”; Biến đo lường thành tố “Lập kế hoạch và thực hiện giải pháp”, hành vi “Sử dụng chiến lược đa dạng để tìm giải pháp, giải thích



Hình 8. Cấu trúc NHCH đánh giá NL GQVĐ

và xác nhận”, mức độ khó là 4.

Ba bài test đánh giá NL GQVĐ của HS cuối tiểu học, cuối trung học cơ sở và cuối trung học phổ thông lần lượt có 36, 36 và 40 câu. Bài test thứ nhất và thứ hai có 13 CH câu, bài test thứ hai và thứ ba có 13 CH câu. Các bài test đã thử nghiệm từ 900 HS ở Vĩnh Phúc, Đắk Lắk và Nghệ An.

2.3. Định cỡ CH. Những CH cầu là cơ sở để cân bằng độ khó các CH và NL các HS ba cấp học. Và tất cả các CH đều được định cỡ theo mô hình một tham số của IRT.

2.4. Đưa CH vào FastTEST

a) Thiết lập cấu trúc ngân hàng gồm ba cấp độ: Cấp học (tiểu học, trung học cơ sở, trung học phổ thông); Các thành tố của NL GQVĐ; Mức độ khó (xem hình 8).

b) Nhập các CH vào từng danh mục con của ngân hàng và đặc điểm tham số CH. Mỗi CH sẽ được gán một mã định danh. Các CH trong test tiểu học sẽ được gán mã TTH01 đến TTH36 và các CH trung học cơ sở, trung học phổ thông tương tự.

Ví dụ: Câu 1 test tiểu học được điền đầy đủ các thông tin trong các mục Identifier, Text, Information, Statistics như *bảng 1*.

Bảng 1. Các thông tin câu 1 trong FastTEST

Unique ID: 8 Item Identifier: TTH01 Description: Thiết bị nghe nhạc Keywords: Công dụng, số học, đại số, th, th	Identifier: Unique ID: do phần mềm tự đánh Item Identifier: Mã định danh CH Description: Theo nhiệm vụ đo lường Keywords: Bối cảnh Toán học, thành tố tìm hiểu VB, mức độ NL 2
Chủ số 2 trong giá của bộ loa máy tính có giá trị là: A. 2 000 đồng B. 20 000 đồng C. 200 000 đồng D. 2 000 000 đồng	Text: Nhập nội dung câu hỏi
Author: Dang Xuan Cuong Source: Date Created: 11/25/2016 Item Response: <input type="radio"/> Multiple Choice <input type="radio"/> Multiple Response <input type="radio"/> True/False <input type="radio"/> Free Response <input type="radio"/> Inventional <input type="radio"/> Survey <input checked="" type="radio"/> Score This Item Answers: Number of Answers: 4 Key(s)/Correct Answer: A B C D E F G H I <input type="radio"/> Numeric <input type="radio"/> Alphanumeric	Information: Điền các thông tin về tác giả, loại CH, phương án trả lời đúng
IRT Model: <input checked="" type="radio"/> 1 - Parameter / Rasch <input type="radio"/> 2 - Parameter <input type="radio"/> 3 - Parameter <input type="radio"/> None IRT Parameters: a: 1.0000 b: -4.2900 c: 0.0000 Maximum item information is 0.723 at theta = -4.290 P-Value: 0.0000 Item Total Correlation: 0.2300 User1: 0.0000 User2: 0.0000	Statistics: Thống số b, P-Value, Item Total Correlation

Order	Item Identifier	Description	Bank Path	Date Added	Exercises	Unique ID
1	TTH03	Thiết bị nghe nhạc	GGVĐ>Toán TH/Lap KH TH	12/2/2016	73	50
2	TTH04	Thiết bị nghe nhạc	GGVĐ>Toán TH/Lap KH TH	12/2/2016	74	13
3	TTH06	Chu kỳ sinh thái	GGVĐ>Toán TH/Lap KH TH	12/2/2016	75	15
4	TTH08	Chu kỳ sinh thái	GGVĐ>Toán TH/Lap KH TH	12/2/2016	76	17
5	TTH29	Giảm giá ở siêu thị	GGVĐ>Toán TH/Lap KH TH	12/2/2016	77	43
6	TTH12	Tàu thủy đến	GGVĐ>Toán TH/Lap KH TH	12/2/2016	78	22
7	TTH17	Thảm quan so thụ	GGVĐ>Toán TH/Lap KH TH	12/2/2016	79	28
8	TTH20	Mua vé máy	GGVĐ>Toán TH/Lap KH TH	12/2/2016	80	32

Hình 9. Các CH của bài test dự kiến

c) Lắp ráp bài test đánh giá NL GQVĐ: lựa chọn các CH trong bài test đo NL HS tiểu học có NL từ mức 2 đến 5 bằng công cụ *Search for items*, kết quả là FastTEST liệt kê 8 CH trong NHCH thỏa mãn (xem hình 9).

Bằng cách kết hợp các phương pháp, kỹ thuật của IRT và thuật toán TNTN của Lê Xuân Tài và Đặng Hoài Phương (2015), chúng tôi đã xây dựng một quy trình phát triển NHCH bao gồm cả 6 hợp phần chính của David J. Weiss (2011).

Trên cơ sở các chức năng sẵn có của phần mềm FsatTEST của ASC, chúng tôi đã xây dựng NHCH đánh giá NL GQVĐ của HS tiểu học, THCS. NHCH này sẽ tiếp tục được thử nghiệm độ chính xác trong thực tiễn, từ đó sẽ được sử dụng trong nhà trường phổ thông. □

Tài liệu tham khảo

[1] Assessment Systems Corporation (2010). *User's manual for the FastTEST 2.0 Item Banking and Test Development System*. St. Paul MN: Author. Available from <http://www.assess.com>

[2] David J. Weiss (2011). *Item Banking, Test Development, and Test Delivery*. In Press, 2011 The APA Handbook on Testing and Assessment Kurt F. Geisinger, Editor-in-Chief Washington DC: American Psychological Association.

[3] Frank B. Baker & Seock-Ho Kim (2004). *Item Response Theory-Parameter Estimation Techniques*. Marcel Dekker

[4] Lê Xuân Tài - Đặng Hoài Phương (2015). *Xây dựng mô hình trắc nghiệm thích nghi trên cơ sở lý thuyết đáp ứng câu hỏi*. Hội nghị Khoa học “Một số vấn đề tiên tiến trong công nghệ thông tin và truyền thông” (SAICT 2015), số 9; tr 5-17. Trường Đại học Bách khoa Đà Nẵng.

[5] Dang Hoai Phuong, Shabalina O.A., Kamaev V.A.. *Proceedings VSTU (2012). Adaptive testing algorithm design methods*. Series “Actual problems of management, computer science and informatics in technical systems”. No: 4. Pages: 107-113. Year 2012.

[6] Nguyễn Thị Lan Phương (2015). *Phương pháp, kỹ thuật xây dựng chuẩn đánh giá năng lực Giải quyết vấn đề, năng lực Đọc hiểu*. NXB Giáo dục Việt Nam.