

MCQ テストの特性

川崎医科大学 数学教室, 薬理学教室*

有田清三郎・斎藤泰一*

(昭和62年9月3日)

Characteristics of Multiple-Choice Questions

Seizaburo ARITA and Taiichi SAITO*

Department of Mathematics, Pharmacology, Kawasaki Medical School,*

Kurashiki, 701-01, Japan

(Received on Sept. 3, 1987)

概 要

多肢選択テスト (Multiple-Choice Questions : MCQ) は国家試験など, 広い分野にわたって採用されている。従来のテスト理論では, MCQ の「あて推量」の効果が取り上げられてきたが, 今回, 「あて推量」とは別の, MCQ 特有の得点に影響を与える因子について検討した。

我々は確実に正答ができる知識の状態に焦点を当て, MCQ 特有の因子(1)正答ターミナル, (2)不用肢, (3)禁止肢を発見した。これら(1)~(3)の存在はこれまで明瞭な形で示されていなかった。(1)は受験者の立場から, (3)は出題者の立場からも重要な因子である。

本稿では, (1)~(3)の因子を2連式MCQテストの具体例で説明し, 数学モデルを使ってその存在を数式で示した。またこれらの因子により, MCQテストでは得点が受験者の知識量に比例しないことが明らかになった。

Abstract

Multiple-Choice Questions (MCQ) has been commonly used for many kinds of qualifying examinations. It has been well known that examinees, who know nothing about, can hit scores by random guessing in the MCQ. In this paper we study the characteristics of MCQ with regard to the factors which can distort scores.

In addition to random guess, we found the following factors:

- 1) If an examinee knows the particular sets of items from given items, he can correctly get the right answer. (Correct Terminals)
- 2) Even if an examinee can judge the correct or false of a particular item, he can not utilize this item for seeking a correct answer. (Unnecessary Items)
- 3) If an examiner sets most difficult items at the particular sets of positions, this test becomes a puzzle. (Forbidden Items)

We discussed above factors using a concrete example of "Doublet Type" of MCQ and showed that the scores of MCQ were not proportional to examinees' knowledge levels.

1. はじめに

多肢選択テスト (Multiple-Choice Questions, 以下 MCQ と略す) は医師国家試験をはじめ, 歯科医師, 看護婦, 臨床検査技師, 栄養士, リハビリテーション技師 (OT, PT) 等の国家試験, 公務員採用資格試験, 共通1次テストなど, 広い分野にわたって実施されている。

MCQ では5肢択一式, 2連式, 3連式, 複合連式, Kタイプなどの種々の形式があるが, 5肢択一式では5つの選択肢が与えられ, その中に1個の正選択肢がかくされている。したがって, 受験者は何も知らなくても偶然に (確率 $1/5$ で) 正解を言い当てることことができる。これが従来から指摘されてきた「あて推量」(厳密には正答率 $1/5$ のあて推量) である。MCQ ではあて推量で得点を増加することができる。

これに対して, 我々はこのあて推量とは別の, MCQ 特有の得点に影響を与える因子を発見した。すなわち

- (1) 与えられた選択肢のうち, 特定の1肢または2肢を知っていれば, 確実に正答ができる。
(正答ターミナル)
- (2) 与えられた選択肢の中には, 正答に寄与しない選択肢が存在する場合がある。(不用肢)
- (3) 与えられた選択肢のうち, 特定の1肢または2肢に難問を設置すれば, 受験者はもはやあて推量なしでは正答できない。(禁止肢)

本稿では, MCQ テストの特性のうち, 得点に影響を与える因子として, 上記(1)~(3)を2連式 MCQ テストの具体例をもとにして説明し, 数学モデルを使って(1)~(3)の存在を示す。またこれらの因子の存在により, MCQ テストでは得点が受験者の知識量を正しく反映していないことを示す。

2. MCQ テストの種類

まず, MCQ テストのかんたんな実例を示そう。

問1

次の地名で岡山県にない町村名はどれですか。

- (1) 宇野
- (2) 内海
- (3) 瀬戸
- (4) 加茂
- (5) 阿波

正解を次の解答コードから選びなさい。

- a (1), (2) b (1), (5) c (2), (3) d (3), (4) e (4), (5)

問1の(1)~(5)は選択肢，a~eは解答コードと呼ばれ，MCQテストは設問，選択肢，解答コードの三要素で構成されている。MCQテストでは，受験者は，正しい選択肢の番号を答えるのではなく，選択肢の組合せで作られた解答コードの記号(a~e)で答える。解答コードは，選択肢を種々組合せて作るので，MCQテストでは種々の形式の問題が使用されている。医師国家試験に用いられているMCQテストの代表的な種類を表1に示す。

表1. MCQテストの種類

問題は，設問，選択肢，解答コードからなる。

設問 △△△について，正しいものは？

- 選択肢**
- (1) _____ .
 (2) _____ .
 (3) _____ .
 (4) _____ .
 (5) _____ .

解答コード a~eの中から正解を1つ選ぶ

1) **単一式**

a.(1) b.(2) c.(3) d.(4) e.(5)

2) **2連式**

a.(1),(2) b.(1),(5) c.(2),(3)
 d.(3),(4) e.(4),(5)

3) **3連式** (2連式の裏)

a.(1),(2),(3) b.(1),(2),(5) c.(1),(4),(5)
 d.(2),(3),(4) e.(3),(4),(5)

4) **複合連式** (選択肢は(1)~(4)の4つ)

a.(1),(3),(4)のみ b.(1),(2)のみ c.(2),(3)のみ
 d.(4)のみ e.(1)~(4)のすべて

5) **単一式変形** (選択肢は(1)~(4)の4つ)

a.(1) b.(2) c.(3) d.(4) e.(1)~(4)のいずれでもない(N)，またはe.(1)~(4)のすべて(P)

3. 2連式MCQテスト

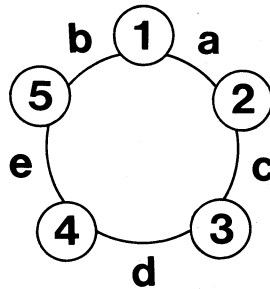
設問に対し正しい選択肢を正選択肢(○)，正しくない選択肢を(×)とする。MCQテストにおけるすべての選択肢は上記の○か×のいずれかに分類される。表1のMCQ形式では，選択肢は5個(または4個)与えられているが，5個の選択肢が与えられている場合には，○と×の組合せで考えると $2^5=32$ 通りの配列がある。表1では解答コードは常にa~eの5組となっているから，MCQテストはこの32通りの配列から，ある規則性を有した5組をとりだして解答コードを構成している。

2連式MCQテスト(以下，2連式と略す)は，解答コードが2個の正選択肢で構成され，解答コードの規則性がわかりやすい形式である。以下，本稿では，この2連式を対象にしてMCQの特性を論ずる。

2連式における選択肢と解答コードの関係は表2の通りである。選択肢(1)~(5)を円周上に並べると，各解答コードは2つの○が連になっていることから，「2連式」と命名された^{1), 2), 3)}。また各選択肢ごとにみると，○の位置がa~eのいずれかの2ヶ所にあり，選択肢ごとの公平性も備っている。

表2 2連式MCQにおける選択肢と解答コードの関係

		選 択 肢				
		(1)	(2)	(3)	(4)	(5)
解 答 コ ー ド	a	○	○	×	×	×
	b	○	×	×	×	○
	c	×	○	○	×	×
	d	×	×	○	○	×
	e	×	×	×	○	○



4. 2連式における正答ターミナル、不用肢・禁止肢

(1) 正答ターミナル

2連式で確実な正答ができるためには、どのような知識が必要であろうか。これを問1の具体例で考えてみよう。問1で与えられた選択肢（地名）について、それぞれが岡山県内にない町村名かどうかをすべて知っていれば、もちろん、確実に正答できる。またこの問題で、設問の岡山県にない町村名（正選択肢：○）を2個知っていても、必ず正答できる。しかしそれ以外の知識でも、確実な正答ができないだろうか。

問2

次の地名のうち、岡山県にない町村名はどれですか。

(1)

(2) 内海

(3) 瀬戸

(4) (5)

a (1), (2) b (1), (5) c (2), (3) d (3), (4) e (4), (5)

選択肢の中の「内海」町は香川県（小豆島）や広島県には在るが、岡山県には存在しないことと「瀬戸」町は愛知県にも在るが、岡山県にも在ることがわかったとしよう。換言すると、(2)が○で、(3)が×の知識を解答コードに適用しただけで、(2)が○なので、aとcだけが残る、さらに(3)が×なのでcが消去され、aがただひとつ残る。すなわち正解はaとわかる。これは

あて推量でなく、確実な正答である。隣接する1個の○と×の部分的知識だけで確実に正答ができる。

また瀬戸町と阿波村が岡山県にあるということを知っていれば、(隣接していない2個の×の知識で)、(3)と(5)に関係した解答コードを消去してゆき、aだけが1つ残る。これも部分的知識だけで、確実に正答できる例である。さらに、阿波村は岡山県にあり(×)、「宇野」は宇野港という地名はあっても、現在では宇野町、宇野村が岡山県にないこと(○)がわかっている場合には、(1)の○と(5)の×の知識だけで、正解aを正答できる。もちろん「宇野」が(○)、「内海」が(○)を知っていれば、その知識だけでそくぞに正答できる。

このような部分的な知識だけで、確実に正答できる知識の組合せを「正答ターミナル」と命名した^{1), 2), 3)}。2連式では、このような正答ターミナルは(○○---), (-○×---), (-×-×), (○---×)の4組である。(2連式の正答ターミナルがこの4組の知識であることを後述の数学モデルで証明する。)

逆の見方をすると、この正答ターミナルは、特定の2肢さえ知っていれば、他の選択肢の知識は全く無くてもよいことを意味している。問1の具体例で考えると、例えば(2)の「内海」(○)の知識と(3)の「瀬戸」(×)の知識だけで正解が見つかり設問に対する正選択肢の内容(第1肢の内容)を知らなくてもよいことになる。

次に、この正答ターミナルがどのような2連の規則性に起因しているかを検討しよう。

いま、解答コードaに注目すると、aの○×配列は○○×××で○群(○が2個)と×群(×が3個)の2群で構成され、かつ○と×がそれぞれ「連」を構成している。○と×の入り乱れた混在配列ではない。解答コードbも選択肢(5)と(1)を循環させて考えるとaの○○×××と同様な配列である。したがって、解答コードa~eの中で、正答コードを見つけるとは、○○×××を○群と×群の2つの群に分離すること、換言すれば○群の先頭(または終端の○)の位置、または×群の先頭(または終端の×)の位置、あるいは○群と×群の境界を見つけることになる。以上のことから、2連式における正答ターミナル(○○---), (-○×---), (-×-×), (○---×)は連の規則性を活用した○○×××配列における○群、×群の分離を可能にする知識である。

(2) 不用肢

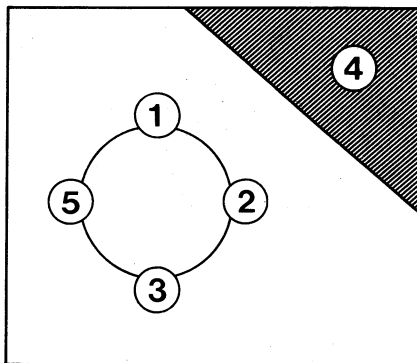
問1の選択肢(4)加茂町が岡山県にあるという知識は、それ単独ではもちろんのこと、他の知識と結合させても、確実な正答(正答ターミナル)にはなりえない。このことは、前述の正答ターミナルのいずれにも、(4)が含まれていないことからわかる。したがって、第4肢は確実な正答を見つけるには不用の選択肢である。(4)の知識はただ単にあて推量の正答率を増やすだけである。(4)が不用肢であることは○○×××の配列で、×群の中間にあり、(4)の知識に他の1肢○か×の知識が附加されたとき(○--×-), (-○-×-)あるいは(-××-), (---××)の知識では正答ターミナルになりえず、○群と×群を分離できないことから

わかる。

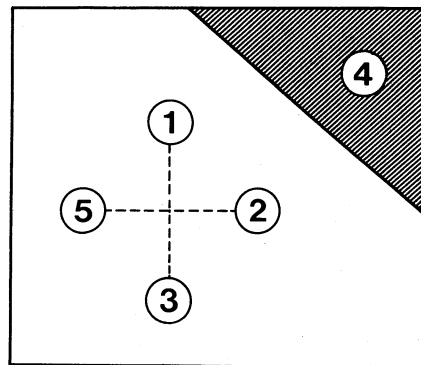
MCQのすべての形式に不用肢が存在するわけではないが、2連式では不用肢が存在し、正解aでは第4肢（一般には×群の中間肢）である。従来、正答に寄与しない選択肢が存在することが、このようにはっきりした形で明らかにはされていなかったが、この不用肢は出題者に「この第4肢に難かしい選択肢（あるいは逆に易しい選択肢）を設置したからといって、問題全体が難かしい問題（易しい問題）になったと解釈してはならない」ことを警鐘している。

解答コードがaのときの正答ターミナルと不用肢の関係を示すため、不用選択肢を別の領域に配置すると、第1肢、2肢、3肢、5肢の4つの選択肢が循環し、正答ターミナルは、隣接する2つの選択肢の組になることがわかる。（図1）

解答者は、正答ターミナルになる2つの選択肢の組合せ（4組もある）を知れば必ず正答できるのに、不用肢1個と、次項に述べる禁止肢にあたる2個の選択肢、計3個の選択肢（例えば、不用肢(4)と(1)と(3)）を知っていても、解答コードは必ず2つ残るので、そのいずれかはあて推量なしには答えられない。すなわち、○-××-または-○-××の知識は5肢のうち3肢（60%）がわかる知識量があっても答えられないのである。これに対して、正答ターミナルでは5肢のうち2肢（40%）の知識で確実に正答できる。これは知識の量と得点が比例しないという重大な欠点をMCQが有していることを如実に示している。



(a) 正答ターミナルと不用肢



(b) 禁止肢

図1 2連式MCQにおける正答ターミナル、不用肢と禁止肢

(3) 禁止肢

正答ターミナルは「確実な正答」ができるための知識に焦点があてられており、これは受験者の立場から、確実に正答できるためには何個の選択肢をどんな位置で知ればよいかという問いかけに対する解答と考えてもよい。

これに対して、今度は「確実な正答」ができなくなる配置すなわち、どの位置の選択肢を何個難かしくすれば、あて推量なしでは正答できなくなるかを検討する。

この問いに対する解答は次の通りである。図1のすべての正答ターミナルが出来なくなるように難肢（受験者の知識レベルを超えた内容の選択肢）を配置せよ。例えば問1の具体例で、(1)だけを難肢にすれば(1, 2), (1, 5)という正答ターミナルは破壊されるが、(2, 3), (3, 5)の正答ターミナルは残存するから、(2), (3), (5)を知っている受験者はあて推量なしで正答できる。そこで(1), (3)の両選択肢を難肢にすると、すべての正答ターミナルは破壊される。受験者は他の(2), (4), (5)の知識を総動員しても、あて推量以外では正答できない。

問3 次の地名のうち、岡山県にない町村名はどれですか。				
(1) ■■■■■		■■■■■は受験者の知識レベルを超えたきわめて難かしい選択肢である。		
○ (2) 内海				
(3) ■■■■■				
× (4) 加茂				
× (5) 阿波				
a (1), (2)	b (1), (5)	c (2), (3)	d (3), (4)	e (4), (5)

(2), (4), (5)の知識を得た受験者は、(1)か(3)のどちらかが○であることはわかって、確実な正答ができないため、aかcの二者択一を迫られる。このような難肢設置は、あて推量を強要するだけで、もはやテストの意味すら失なわせるため、「禁止肢」と命名した。2連式の禁止肢は(1, 3)と(2, 5)の2組である。

この禁止肢の存在も明確にされていなかった。禁止肢は、文字通り出題者に「MCQテストをあてものにさせないための難肢配列の禁止則」を警報するものである。これによって出題者は不用意に難肢を設置できなくなるであろう。受験者がMCQテストでよく遭遇した二者択一のジレンマは、この禁止肢に起因していたのではないだろうか。

禁止肢を○○×××の配列から考察する。禁止肢は○群と×群をくっきりと分離できなくする難肢の配置である。2連式での禁止肢は、■○■××, ○■××■で○群と×群の境界をわからなくするため、○と×の間に交差させて、難肢を配置させている。この場合にも不用肢(4)は何ら関与していないことがわかる。もちろん、全肢を難肢にしてしまえば、あて推量なしでは正答できないから、禁止肢を難肢の最小限の配置という意味でつかっている。

5. 数学モデルによる検討

2連式の正答ターミナル、不用肢、禁止肢を数学モデルを使って検討する。2連式では5つの選択肢が与えられている。いま、各選択肢の理解を「わかる (known)」と「わからない (unknown)」のいずれかと考える。5つの選択肢では $2^5=32$ 通りの知識の組合せがある。このうち「わかる」には正選択肢を理解できる場合 (○) と、不正選択肢を理解できる場合 (×) がある。また第 j 肢の選択肢のわかる確率を p_j ($0 \leq p_j \leq 1$; $j = 1, 2, \dots, 5$) とする。

2連式で正解が a の場合について議論する。(他の解答コードの場合についても、2連式の規則性によって、この議論は一般性を失わない。) $2^5=32$ 通りの知識の状態のそれぞれに対応させて、それが生起する確率と正答率を求め、このうち、正答率が1となる生起確率の和を P とすると、 P は一般に $(p_1, p_2, p_3, p_4, p_5)$ の関数で表される。2連式での P は次式で与えられる。

$$P = \{p_1 + (1-p_1)p_3\} \cdot \{p_2 + (1-p_2)p_5\} \quad (1)$$

1. 正答ターミナル

正答ターミナルは、確実な正答ができる知識の状態で、それらは $P=1$ の解で与えられ、 $p_1=p_2=1, p_1=p_5=1, p_2=p_3=1, p_3=p_5=1$ となる。すなわち2連式の正答ターミナルは (1, 2), (1, 5), (2, 3), (3, 5) の4組である。

2. 不用肢

P は一般に、 $(p_1, p_2, p_3, p_4, p_5)$ の関数で与えられるが、2連式の P には p_4 が含まれていない。すなわち第4肢は確実な正答に寄与しない選択肢である。以上より2連式には不用肢が存在し、それは第4肢である。

3. 禁止肢

あて推量なしでは正答できない、すなわち、確実な正答ができなくなる禁止肢は、 $P=0$ の解で与えられ、 $p_1=p_3=0, p_2=p_5=0$ となる。すなわち2連式における禁止肢は (1, 3), (2, 5) の2組である。

2連式の正答ターミナル、不用肢、禁止肢は知識の組合せと、解答コードとの規則性によっても得られるけれども、数学モデルを使用することによって、論理的に導出できる。

6. 得点に影響を及ぼす因子

MCQ の特徴のひとつは、MCQ の形式 (正選択肢の個数や組合せ) のいかににかかわらず、正解コードが一つであって、一つでしかない点である。この特徴による「あて推量」が従来から議論され、あて推量もとりわけ正答率 $1/5$ のあて推量だけが問題にされてきた。しかしながら、MCQ の形式に由来する種々の因子は具体的には検討されていない⁶⁾。

本稿では、受験者の知識を誤りの知識なしに限定して、「確実に正答できる」知識を検討した。MCQ では、2つの知識でもそれが正答ターミナルであれば、確率1で正答でき、逆に3つの知識でも、正答ターミナルでなければ、あて推量以外では正答できない。また正答ターミナル

を含んでいる知識では、知識の量が多くても得点は同じで、高くなることも、低くなることもない。以上のことから、MCQでは「知識の量が多ければ、多いほど得点が高い」という説は迷信であり^{7), 8), 9), 10)}、知識の量と得点が比例しないことが示された。

本稿では、誤りの知識なしで知識と得点の関係を考察したが、現実には、思い違いなどの誤りの知識があるため、知識から得点への経路には、正答ターミナル、あて推量のほかに誤答ターミナル（必ず誤答してしまう知識）があり、誤りの知識はMCQで特異な働きをする^{5), 7), 8)}。すなわち、1個の誤りの知識が誤答ターミナルに導びく場合、誤りの知識が多くなると、知識が解答コードの組合せと矛盾を生じ、どこかの選択肢の知識がまちがっていることに気づき、誤りの知識を放棄し、あて推量により正答できる場合がある。このとき、誤りの知識が多くなると、誤答ターミナルから離脱し、あて推量で正答領域へ移行したことを意味する。このことから、知識の量と得点が比例しないことが示される。

MCQのもうひとつの欠点は、得点から受験者の知識量や実力が推測できないことである。MCQでは、知識から得点に到るには、多数の経路がある。誤りの知識なしで、選択肢の知識を「わかる」、「わからない」の2種類に限定しても $2^5=32$ 通りの経路がある。MCQで正答した受験者は、5肢のすべてを知っていたのか、不十分な知識でも正答ターミナルに達して確実に正答したのか、不十分な知識で正答ターミナルに達せず、あて推量（正答率 $1/2$, $1/3$ 等）により正答したのか、知識なしのあて推量（正答率 $1/5$ ）による正答だったのか、採点者には全く不明である。ただ受験者のみを知るだけである。32通りの経路があっても、正答ターミナル以外のあて推量による正答の割合はきわめて少ないであろうという暗黙の了解があるように思われるが、我々はこれらの知識の生起確率と得点分布及びシミュレーション結果を別に求めている^{3), 10)}。

MCQテストは「知識の有無と正答、誤答とが正しく対応するテスト」ではない。換言すれば「受験者の知識があれば、必ず正答し、知識がなければ、必ず誤答する」ことが明確に示されず、また逆に「正答した受験者は知識があったこと」が、「誤答した受験者は知識がなかったこと」もはっきりわかるテストではない。MCQテストは知識の量を得点に正しく反映できない、信頼性の低いテストであり、国家試験や資格試験などにMCQを使用すべきではない。

文 献

- 1) 斎藤泰一・有田清三郎・那須郁夫 (1981) : 多肢選択問題は果たして客観的評価法といえるか, 医学教育13, 251-255
- 2) 有田清三郎・斎藤泰一・那須郁夫 (1982) : 多肢選択テストの部分的知識による得点増加を評価するための数学モデル, 行動計量学会, 10, 53-66
- 3) 那須郁夫・有田清三郎・斎藤泰一 (1983) : 多肢選択テストの形式が受験者の得点に及ぼす影響——受験モデルによるコンピュータシミュレーション, 医学教育14, 410-418
- 4) 斎藤泰一・有田清三郎・那須郁夫 (1983) : 現行医師・歯科医師国家試験問題の形式は能力評価に適さない, 医学教育14, 405-409

- 5) 有田清三郎・斎藤泰一・那須郁夫 (1984) : 多肢選択テストにおいて受験者の誤った知識が得点に及ぼす影響, 医学教育15, 51-53
- 6) 日本医学教育学会学部教育委員会編 (1982) : 医学教育マニュアル4, 評価と試験, 篠原出版, pp68-93
- 7) 斎藤泰一・有田清三郎・那須郁夫 (1984) : 多肢選択テストが医学教育に及ぼす影響, 医学教育振興財団助成報告書
- 8) 斎藤泰一・有田清三郎・那須郁夫 (1985) : 多肢選択テストにおける出題者の錯覚, 医学のあゆみ, 135, 107-110
- 9) Taiichi Saito, Seizaburo Arita and Ikuo Nasu (1986) : Untoward effects on the evaluation of examinees caused by the format of Multiple-Choice Questions, Proc. of the 2nd Japan-China Symposium on Statistics : 220-223
- 10) Seizaburo Arita, Taiichi Saito and Ikuo Nasu (1986) : Score Distribution of Examinees with Partial Knowledge by the Test of Multiple-Choice Questions, Proc. of the 2nd Japan-China Symposium on Statistics : 5-8