

## 多肢選択問題における当て推量の現状とその対処法について

斎藤 泰一, 笠井 裕, 有田清三郎\*, 那須 郁夫\*\*

多肢選択問題 (MCQ) には当て推量による解答があることが問題になっている。しかしこれまで具体的にどの問題が当て推量で答えたのかを判定した報告は見当たらない。MCQ の出題時に、なぜその選択肢を選んだかという理由を書かせてみたところ、正答者の中にも高率に理由の間違いや白紙答案がみられた。また誤答した者の中にも理由は正しく書いている者も相当みられた。このように、MCQ は知識の程度を正しく評価し得るような問題形式ではないことが明らかになった。筆記や口頭試問の方がより正しく評価し得るものと考えられる。

(平成元年4月18日採用)

### Random Guessing in the Multiple-Choice Question Tests: Status in Quo and Its Exclusion

Taiichi Saito, Yutaka Kasai, Seizaburo Arita\* and Ikuo Nasu\*\*

In the multiple-choice question (MCQ) examinees can answer by random guessing. The examiner, however, can not point out which question was answered by random guessing. Then, we asked to write the reason why the examinees selected that answer code. In the correctly answered group more than half examinees could not write the suitable reason. Moreover, some examinees who selected wrong answer wrote the correct reason. There are much more answers guessed randomly than expected by probability model. Therefore, the MCQ test can not estimate the examinee's knowledge properly. (Accepted on April 18, 1989)

*Kawasaki Igakkaishi 15(2): 364-369, 1989*

**Key Words** ① Multiple-choice questions ② Random guessing  
③ Medical education

#### はじめに

医師、歯科医師、看護婦、薬剤師などの国家試験をはじめ、共通一次試験その他小学校から大学までの各種試験、資格試験などに広く用いられている多肢選択問題 (multiple-choice

question, MCQ) にはいくつかの形式があるが、いずれにしても解答が準備されており、その中から選択して正誤を答える形式が一般に用いられている。そのため、正答した者の中にも当て推量によるものがあることが指摘されている。

川崎医科大学 薬理学教室  
〒701-01 倉敷市松島 577

\* 同 数学教室

\*\* 日本大学松戸歯学部 衛生学教室

Department of Pharmacology, Kawasaki Medical School:  
577 Matsushima, Kurashiki, Okayama, 701-01 Japan  
Department of Mathematics

Department of Dental Public Health, Nihon University  
School of Dentistry at Matsudo

五つの選択肢から一つの正解を選ぶ単純五肢択一式（一連式）では、何も知らない人が答えても5分の1は当て推量で正答できると一般には考えられている。また医師国家試験に用いられているような二連、三連、複合連式<sup>1),2)</sup>になると、その形式によって単純に当て推量が5分の1というわけにはいかない。我々は単一式でも単なるランダム・ゲッティングではなしに、受験者がわかる選択肢の数によって得点に変化することを明らかにした。<sup>3)</sup> また二連、三連、複合連式も含めて正答ターミナルという概念を導入してMCQの得点を求める数学モデルを考案し、<sup>4)</sup> さらにコンピュータシミュレーションによる実験をも発表した。<sup>5)</sup> なお三谷も医師国家試験問題の形式による得点の違いを正答肢の位置を中心にして論じている。<sup>6),7)</sup>

いずれにしても、これまでのMCQの得点分析は確率モデルなどを適用した統計的処理であって、個々の解答者が実際にどの選択肢を知っており、あるいは知らなかったかということは全く推量の域を出なかった。したがって本当に当て推量をしたかどうかとも全く不明のまま

推論を重ねていたにすぎない。そこで我々は実際の問題で受験者がどのようなルートで解答するのかを明らかにする目的で本研究を行い、またこの方式を使用すると当て推量による得点増加を除外できるかもしれないと考えたので報告する。

### 方法と結果

#### I. 解答ルートを明示させる方法

三年生の薬理学の講義が始まって間もないときに、学生にテストの目的はMCQの解答経路を調べることであり、薬理学の点数とは無関係であることを説明した上で試験を実施した。形式は、心筋梗塞の患者がワルファリンを服用していたが、肩の痛みのために整形外科でインドメタシンを投与され、両者の服用を続けているうちに中枢神経系の出血を来した症例を呈示し、解答として選択肢を二つ、すなわち、

- (1) ワルファリンの血中遊離濃度が上昇した
- (2) " " " " 低下した

のいずれが正解かを選ばせた。さらにTable 1のような組合せを示し、自分はこの経路でその解答を選んだかを書かせ、またその解答を選んだ理由を書かせた。124名のうち、正答した者は49名(39.5%)しかいなかった。正答者、誤答者とも、そこに至った経路をTable 1の2)の項目に記入させて調べてみた結果をTable 2に示した。

Table 1. Questionnaire

1) どちらの選択肢が正しいか？

選択肢	(1)	(2)
正答	○	×
誤答	×	○

2) 解答に到達した経路を書きなさい

○：正しい    ×：誤り  
 ?：よくわからない

選択肢	(1)	(2)	
a	○	×	} →正答
b	○	?	
c	?	×	
d	×	○	} →誤答
e	?	○	
f	×	?	
g	?	?	} →当て推量

3) その選択肢を選んだ理由を書きなさい

Table 2. Analysis of answer routes

	解答の組合せ		人数	%
	(1)	(2)		
正答者	a	○ ×	7	14
	b	○ ?	14	29
	c	? ×	9	18
	g	? ?	19	39
	計		49	100
誤答者	d	× ○	18	24
	e	? ○	16	22
	f	× ?	14	19
	g	? ?	26	35
	計		74	100

正答した者49名のうち、(1)が正しく、(2)が誤りであることがわかって答えたもの(a)は7名(14%)にすぎなかった。しかし、(b)、(c)につけた者は23名(47%)あり、これらはいずれも正答になるルートであり、これらの三つのグループに属するものは61%であった。これに対して(g)と答えた者は、どちらが正しいかわからずに、当て推量で正答した者であり、19名(39%)いた。すなわち正答した者のうちの1/3以上が当て推量で正答したことになる。

このように、自分が解答した判断の経路を書かせると、どのように信じて解答したかということがわかることになる。しかし、実際の試験になると正直に書く人はいなくなると思われるので、この方法は実用的ではない。そこで次のステップとしてなぜそうなるのかという理由を書かせた。その結果をTable 3に示した。

正答者49名のうち、理由が正しく書けていた者は23名(47%)であったが、理由が正しいことと、選択肢の選び方との間には大きな矛盾がみられた。この問題は、薬理学の講義が始まって2回目に行ったものであり、かつ目的が成績の判定ではなかったので難度の高いものを用いたことが影響したのであろうと思われるが、選択肢(1)(2)を正しく(O, X)とした者7名のうち、理由が正しかった者は3名、誤りであった者も3名、理由のわからなかった者(多分当て推量でそう思ったと考えられる)1名であった。正答者49名のうち、理由が誤りであった者が20名(41%)もいたことは、このようなO×式の問題形式は客観試験と呼ばれて、誰が点数をつけても採点者の主観が入ることがないのでよいテストだと一般に信じられているが、知識の状態と得点の間にはこのような矛盾が存在することが明らかになった。しかも当て推量も、従来行われている確率計算とは異なった頻度になることもわかった。すなわち受験者の能力と、問題の質によって思わぬ割合で当て推量が行われるのであり、統計学的な処

Table 3. Reason of the choices in the correctly answered group

選 択 肢		解 答 者 数	理 由 記 載 欄		
(1)	(2)		正しい	誤り	白紙
○	×	7	3	3	1
○	?	14	9	5	0
?	×	9	5	4	0
?	?	19	6	8	5
正 答 者 計		49名	23名 (47%)	20名 (41%)	6名 (12%)

理をしても個々の例については何の示唆も与えるものではない。

## II. 多肢選択テストで理由を書かせる方法

四年生の利尿剤の講義が終わった時期にテストを試みた。その中の一問は「Spironolactoneを投与すると、尿中のKの排泄が増加する」という設問で、これが○か×か、およびそれはなぜかということを書かせた。受験した者140名中、正答(×)した者は116名(83%)であった。説明の項をみると、aldosteroneに拮抗してKの排泄が減少すると書いた者(これは正しい)は88名で、正答者の中の76%であった。22名はaldosteroneのことは記載しておらず、その中の10名はKの排泄は変化しないと書いていた(これは誤り)。しかしこの10名が選択肢に×をつけていたことは正しかったのである。別の3名はADHに拮抗すると書いており(これも誤り)、正答(×)しているにもかかわらず説明には逆のことを書いているものが1名いた。1名はNaの排泄が増加すると書いていたが(これは正しい)、このことは出題者が意図したKの排泄とは無関係な答えになったが、このような反応はMCQで○×だけをつけさせる形式では考えられないことである。

## III. 問題の形式による違い

四年生の循環器系の系統別講義の試験に一連式(単純五肢択一式)2問と、複合連式(Table 4)で正答の解答コードがaである場合とcである場合の問題を出して、それぞれ正答の解答コードを選ばせた上、その理由を書かせた。複

**Table 4.** Complex type of MCQ

設 問：安静時狭心症の発作の予防に有効な薬物は どれか
選択肢：(1) ニトログリセリン錠 (2) 硝酸イソソルビドテープ (3) ニフェジピン (4) 亜硝酸アミル
解答コード
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

**Table 5.** Relation between the formats of questions and answers

		第1問 一連式	第2問 複合連 (a)	第3問 複合連 (c)	第4問 一連式
正 答 者		27 (19)	39 (28)	25 (18)	69 (50)
理 由	正 誤	(56)	(44)	(72)	(39)
	白 紙	(9)	(16)	(12)	(42)
		(11)	(15)	(16)	(19)
誤答者で理由 が正しい者		1 (1)	28 (20)	7 (5)	0 (0)

( ) = %

合連式の問題は、二連、三連式と同様に、医師や歯科医師の国家試験に常に用いられている形式である。またこれらの形式の問題は、特定の選択肢二つがわかれば、残りの選択肢が空白になっていても必ず正答できるもので、<sup>2),8)</sup>我々はこの正答に至る選択肢の組合せを「正答ターミナル」と名付けた。<sup>4)</sup>すなわち、すべての選択肢に関する正しい知識がなくても、特定の組合せになる二つだけわかれば必ず正答できるし、その組合せの数が、二連式、三連式では4組もある。複合連式では正答の解答コードの位置によって異なってくるので、aでは2組、b, cでは4組、dは5組、eは2組(ただし1組は1, 2, 3, の三つの選択肢が判らなければ解答できない)となっている。こうしてみると、MCQは形式によっては大変易しい問題であるとも考えられるが、誤判断や、一部の誤った知識の下で結論を出すと、知識の割合に応じ

た点数が得られなくて、必ず0点になってしまうのである。<sup>9)</sup>

このテストで内科系の教員と一緒に50題のMCQを出題し、60点以上の得点があったのは139名中38名にすぎなかった。薬理学では4問出題して、理由も書かせた。その結果を**Table 5**に示した。

正答した者の割合は問題によって大きな差がみられたが、これは形式によるというのではなく、問題の質によるものと考えた方がよい。しかし同じ一連式でも第4問のように正答した者の中に理由が間違っている者が42%、白紙が19%もあるということは、わけがわからなくても何となく正答することができる場合がいかにも多いかを示している。逆にみると、正答した者が本当に理解しているかどうかはMCQでは判定できないことになる。また理由が正しいのに誤答している者も複合連式ではかなりみられる(第2問では20%もいた)ことは、正答に無関係な選択肢でもその正誤の判定を誤れば、足を引っ張られて誤答してしまうことを示している。

### 考 察

MCQによるテストでは、受験者が当て推量で答えたのか、本当に知っていて答えたのかは試験官はただ推測するほかはない。選択肢の選び方が正しいか、そうでないかを機械がマークシートで判定できるので、時間が節約できというのが唯一の長所である。試験官の主観が採点に影響しないということも大きな長所であると考えられる人が多いが、試験というのは試験官の出題した問題の○×を書かせるのが目的ではない。いかによく理解しているか、またどんな考え方を持っているのか、どれくらいの可能性を秘めた人間であるかを判定するのも入学試験その他では判定の大きな目的になると考えてよい。これには試験官と受験者の間でコミュニケーションがなければならないはずである。そうすると試験官の個性、主観というものもテストにとっては大切な要素になってくる。すなわ

ち人間の能力の判定には試験官の資質が大きくかかわってくるものであるから、むしろ主観性が入らないことは欠点と考えてもよいはずである。

当て推量で答えたかどうかは、今回のテストのように「なぜその選択肢を選んだのか」を書かせると判断することができることがわかった。しかし、それでもなお理由が正しくても誤答する者がかなりあったりするので、当て推量だけが影響しているのではないと考えられる。誤答した者は、たとえ正しい理由がわかっていたとしても、選択肢の中によく知らないものがあったり、理由にまだ確信が持てなかったりした場合には別の解答コードを選んでしまう可能性が大きいことを示している。

誤答した人は、すべての選択紙を確実に知っていたかといえるが、正答した人が知っていたかどうかはMCQの○×による解答だけでは判定できないことが確かになった。この原因としては、MCQではいつでも答えがあり、その中のどれかに印をつければよいという形式が第一の要素として考えられる。もし筆答試験であれば自分の考えを書くしかないので、でたらめを書けばでたらめであると判定することが可能である。さらに口頭試問を行えば、どの点が間違っているのか、またどんな思考過程でそう考えたのかということも、その場に即して直ちに判定できるので、受験者の反応を十分に観察することができるのである。

MCQで理由を書かせても、正答者で理由が間違っていた人と、誤答者で理由が正しかった人との採点の仕方はどのようにすればよいのかといった問題が残る。

MCQの形式に関しては、一連式(単純五肢択一式)が他の影響が少なくてよいように考えられているが、今回のテストでは問題によって正答者69名うち、間違った理由を書いた者が42名もおり、正しい理由を書いた39名よりも多かったことは(残り19名は理由を書かなかった)、一連式も全く信頼できないことを明らかに示している。

石井ら<sup>10)</sup>は次のような実験を行った。すな

わち選択肢を118問作って1つずつ正誤を答えさせ、次にその組合せを作ってMCQとして25問(一連式5問、二連式7問、三連式6問、複合連式7問)作成した。これによると、選択肢はすべて同一であるにもかかわらず、個々の選択肢の判断とそれを組み合わせた形式とでは正答する割合が異なっており、一つの選択肢のときに正答したものでも組み合わせると正答できなくなることを示している。一肢だけで正しいと思っても、組合せになると確信が持てなくなって誤答してしまうと述べている。この結果は我々の今回の調査でも同じことであるが、別の見方をすれば、一つの選択肢の判断そのものが正しかったかどうかの根拠は何もないことになる。偶然正答したのかも知れない。その確率は1/2である。

石井らの論文によると、一連式と二連、三連、複合連の方式とでは平均点にはほとんど差がないが、分布に大きな差がみられる。すなわち一連式では平均点は77.46点であり、その分布範囲は64点から92点の間に分布しており、上下の差は28点であった。しかるに多連の形式にすると、平均点は78.96点で有意差はなかったが、分布範囲は48点から100点までの52点の広がりを見せていた。石井らはこの理由として受験者の確信度をあげているが、確信度というのは換言すれば知識が確実かどうかということである。これだけの考え方では、多連式になったときはなぜ分布が広がるのかの説明は難しいと思われる。

我々は正答ターミナルの考えで説明している。すなわち一連式では正答ターミナルは2組しかない。これは○の選択肢一つを知っている場合と、×の選択肢四つを知っている場合とである。したがって、この場合は○の選択肢をかなり十分に知らないかと判断しにくいことになる。それに比べて多連になると、二つの選択肢1組で正答ターミナルになり、しかもそれが4組またはそれ以上もあるので、一部の知識で確実に正答できるため、100点までとれるようになるのである。また誤った知識や不確実な知識ではその問題は0点になるか、また二つの

選択コードの判断にまで達したときにわからな  
いため1/2の確率で当て推量をしてしまうか  
ら、得点がどんと下がるか、逆に知らなくても  
1点もらえるようになるためである。このよう  
に、一連でもその他の連でも、MCQの形式は  
受験者の知識の状態を正しく評価できるもの  
ではないので、評価には使用すべきではないと考  
えている。また理由を書かせても、正しい知識  
を持っているかどうかということと、MCQの

○×の結果とは全く矛盾した結果になる場合  
があることがわかったので、理由を書かせてもど  
のように点数をつけるかということには大きな  
困難が存在する。筆記試験や口頭試問を行うの  
がよい方法だと考えられる。

本研究の一部は文部省科学研究費の援助(No.  
62510149)ならびに操風会教育奨励金の援助を受け  
た。

## 文 献

- 1) 齋藤泰一, 有田清三郎, 那須郁夫: 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, 1983
- 2) 齋藤泰一, 有田清三郎, 那須郁夫: 多肢選択問題における出題者の錯覚. 医学のあゆみ 135: 107-110, 1985
- 3) 齋藤泰一, 有田清三郎: 多肢選択問題の正答率に及ぼす不十分な知識の影響. 医学教育 12: 356-360, 1981
- 4) 有田清三郎, 齋藤泰一, 那須郁夫: 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 10: 53-66, 1982
- 5) 那須郁夫, 有田清三郎, 齋藤泰一: 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, 1983
- 6) 三谷一裕: 多肢選択問題への新しいアプローチ「標準正答率」の提案とその応用—(その1). 医学教育 13: 238-250, 1982
- 7) 三谷一裕: 多肢選択問題への新しいアプローチ「標準正答率」の提案とその応用—(その2). 医学教育 13: 399-406, 1982
- 8) 齋藤泰一: 医師国家試験の問題点. 医師国試対策 8: 61-67, 1987
- 9) 有田清三郎, 齋藤泰一, 那須郁夫: 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, 1984
- 10) 石井哲夫, 山本信和: 同一内容による単純真偽と多肢選択形式の比較. 医学教育 19: 14-17, 1988