

項目特性曲線による試験問題の評価と試験成績の関連について —教育改革戦略の礎—

三笠洋明（徳島大学医学部教育支援センター）

赤池雅史（徳島大学医学部教育支援センター）

西村明儒（徳島大学医学部教育支援センター）

1. 目的

知識領域における学修到達度評価方法としては筆記試験を用いることが多い。注意深く作成された多肢選択問題(MCQ)は想起レベルのみならず応用のレベルの知識まで評価できると考えられておりCBT、医師国家試験等で広く用いられている。しかしながら、MCQ を含め通常の試験では学生の学修到達度と試験問題の難易度の区別は容易ではない。そこで各科目の成績と各項目（試験問題）の項目特性曲線、正答率、点双列相関係数の関連を検討することでより良い試験問題の有り方とその改善を通じて教育改善の方略について検討する。

2. 対象と方法

対象者は、20××年の×月から△月に実施した21科目の試験を全て受験した112名。試験形式はマークシートを用いた多肢選択問題(MCQ)である。全科目の得点平均値で5群に分けた。各項目（試験問題総数1334）について5つの成績群における正答率を縦軸に取り、項目特性曲線を描いた。

3. 結果

表1に得点、正答率に関する科目別の記述統計を示した。科目は得点の平均値の降順とした。得点の平均値は46.0から84.7と1.8倍の開きが認められた。

全科目の項目の正答率と点双列相関係数の散布図を図1に示した。正答率は1直下に集積が認められた。中央値は0.848、平均値は0.734。点双列相関係数の範囲は最小値-0.217から最大値0.553。平均値は0.131標準偏差は0.125であった。識別力を有すとされる点双列相関係数0.2以上の項目の割合は34.9%。また50%以上の正答率の割合は78.6%であっ

た。

表1 科目別得点と正答率

科目	得点			正答率	
	最小値	最大値	平均値	平均値	度数
1	67.1	94.1	84.7	0.855	83
2	66.0	96.0	83.9	0.839	50
3	69.9	93.2	83.8	0.845	69
4	68.0	98.0	82.8	0.828	50
5	65.0	94.0	81.8	0.828	60
6	60.0	99.0	81.8	0.779	49
7	58.0	93.0	81.3	0.768	77
8	60.7	96.4	80.8	0.808	56
9	60.0	94.0	78.5	0.785	50
10	56.0	92.0	77.9	0.778	50
11	50.0	92.0	76.2	0.762	50
12	60.0	87.5	75.2	0.752	80
13	54.0	88.0	74.4	0.744	50
14	56.9	89.7	74.2	0.742	58
15	54.7	92.5	73.9	0.739	53
16	52.8	86.1	71.7	0.717	36
17	51.0	81.0	70.2	0.702	100
18	51.7	81.0	69.2	0.692	58
19	46.0	82.0	64.7	0.647	100
20	36.0	78.0	54.3	0.543	50
21	26.5	62.2	46.0	0.460	98
合計	60.7	83.1	74.6	0.734	1327

科目の数字は得点平均値の順位

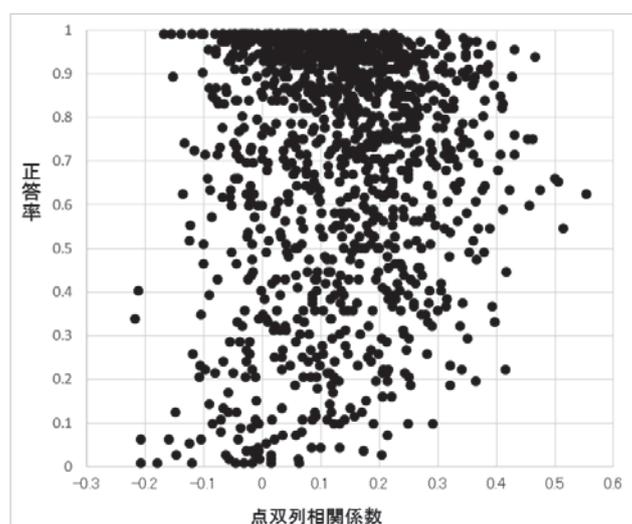


図1 全科目全項目の正答率と点双列相関係数

次に、各科目毎の正答率(50%以上を高)と判別力の指標である点双列相関係数(0.2以上を判別力有)に関し集計を行った結果を表2に示した。平均得点が最も高い科目1では正答率が50%を超える項目の割合は92.8%、得点が最も低い科目21では49.0%と1.9倍の開きが認められた。識別力を有す項目の割合に関しては、最小値は13.9%最大値は48.1%と3倍以上の開きが認められた。また、得点上位科目では40%程度であるが、得点下位科目では10%～20%台が散見され弱い低下傾向が認められた。

表2 科目毎の識別力の有無と正答率の高低による項目の割合

科目	識別力有	正答率高	正答率高		正答率低	
			識別力有	識別力無	識別力有	識別力無
1	43.4	92.8	42.2	50.6	1.2	6.0
2	36.0	88.0	32.0	56.0	4.0	8.0
3	33.3	87.0	33.3	53.6	0.0	13.0
4	40.0	90.0	34.0	56.0	6.0	4.0
5	30.0	88.3	28.3	60.0	1.7	10.0
6	34.7	85.7	28.6	57.1	6.1	8.2
7	48.1	89.6	44.2	45.5	3.9	6.5
8	30.4	89.3	30.4	58.9	0.0	10.7
9	42.0	90.0	40.0	50.0	2.0	8.0
10	36.0	82.0	30.0	52.0	6.0	12.0
11	40.0	84.0	34.0	50.0	6.0	10.0
12	33.8	78.8	23.8	55.0	10.0	11.3
13	38.0	82.0	32.0	50.0	6.0	12.0
14	34.5	81.0	29.3	51.7	5.2	13.8
15	39.6	77.4	26.4	50.9	13.2	9.4
16	13.9	77.8	13.9	63.9	0.0	22.2
17	35.0	69.0	30.0	39.0	5.0	26.0
18	32.8	74.1	27.6	46.6	5.2	20.7
19	23.0	69.0	17.0	52.0	6.0	25.0
20	36.0	52.0	20.0	32.0	16.0	32.0
21	31.6	49.0	18.4	30.6	13.3	37.8
合計	34.9	78.6	29.2	49.4	5.7	15.7

正答率の高低と識別力の有無によるクロス集計(表2)によると、正答率の高い項目群の中の識別力の有る項目の割合が得点下位の科目19～21では17.0～20.0%と小さくなっていった。一方、正答率の低い項目群の中の識別力の無い項目の割合は増加傾向が認められた。特に科目16以降では識別力の無い項目の割合が22.2%～36.0%に上っていた。

4. 考察

試験の特性を正答率と識別力の有無で検討した。学修到達度の高低を識別する力を持つ項目の割合は全体で3割であり、試験の大部分は学修到達度ではない何かを測定している可能性が認められた。識別力を有す項目は正答率の高低に関わらず学修到達度の評価が可能であることから識別力を有す項目の割合を出来るだけ増加させるために適切な問題作成に関する情報提供を行っていく必要があると考え

られた。また、得点下位の科目では、正答率が低かつ識別力の無い項目(曲線例1)の割合が上位の科目に比べて高く22～37%含まれていた。このような項目は学修到達度を測定することは出来ないだけでなく、試験時間を浪費させ、学生の学修意欲をそぐと考えられることから試験の質のみならず教育の質を低下させると考えられた。また、学生はこのような項目に接すると、教科書を理解するという学修を放棄し過去問の答をパターン化して暗記するという逃避行動に走らせる可能性も考えられることから、試験問題の適正化が学生の学修行動の改善に繋がると考えられた。また正答率の低さは講義における適切な情報伝達が出来ていない可能性も含め今後のFDなどを通じ検討する必要があると考えられた。

そういう意味では正答率が高く識別力のある項目(曲線例4)が望ましいとかがえられる。一方、正答率が高く識別力が無い項目(曲線例5)に関しては出題者がそれを必修の事項として出題した場合には適正であると考えられる。しかしながら学修到達度の評価事項として出題した場合には何らかの作問上の設定に瑕疵があったと考えるべきであろう。

5. 考察

試験得点が低い科目はそれ以外の科目に比べ正答率が低かつ識別力の無い項目の割合が高い事が明らかになった。今後、適切な難易度を持ち識別力の高い問題作成のための情報提供を行う事で学生の学習行動への影響を注視していきたい。

