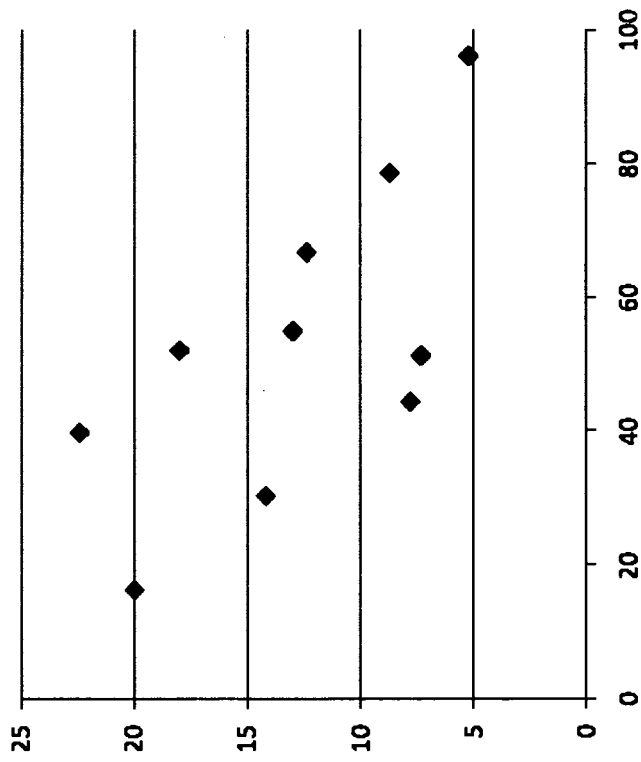


図2 学生と医師の得点、選択項目の平均値における相関

得点と選択問診項目数

$$r = -0.67$$



得点と選択検査項目数

$$r = -0.80$$

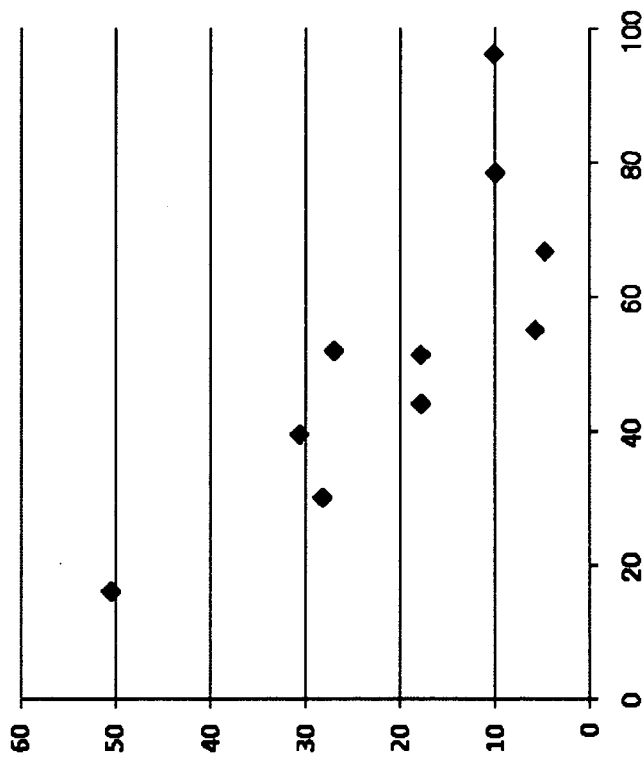
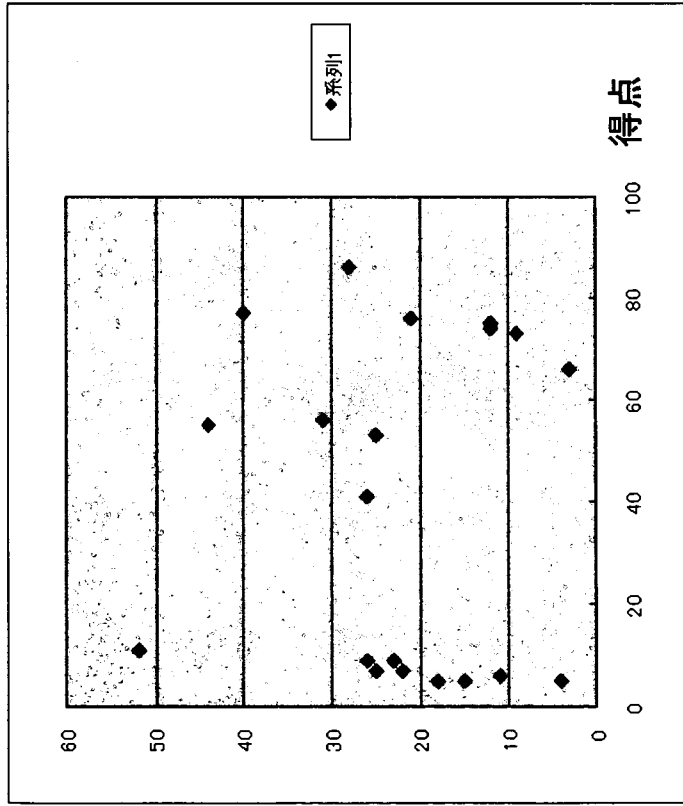


図3 インフルエンザ問題の個々の学生の得点と選択数

得点と選択問診数



得点と選択検査数

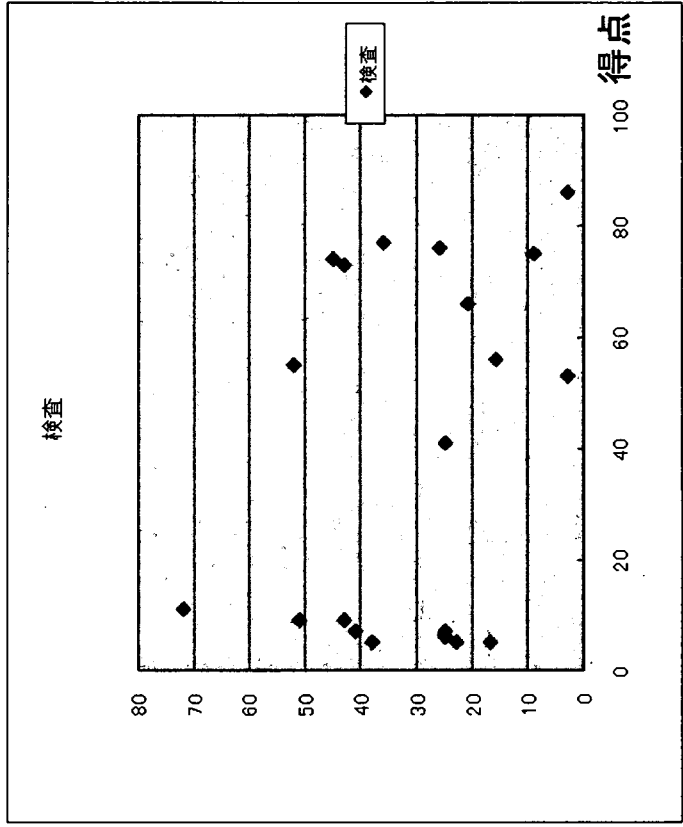
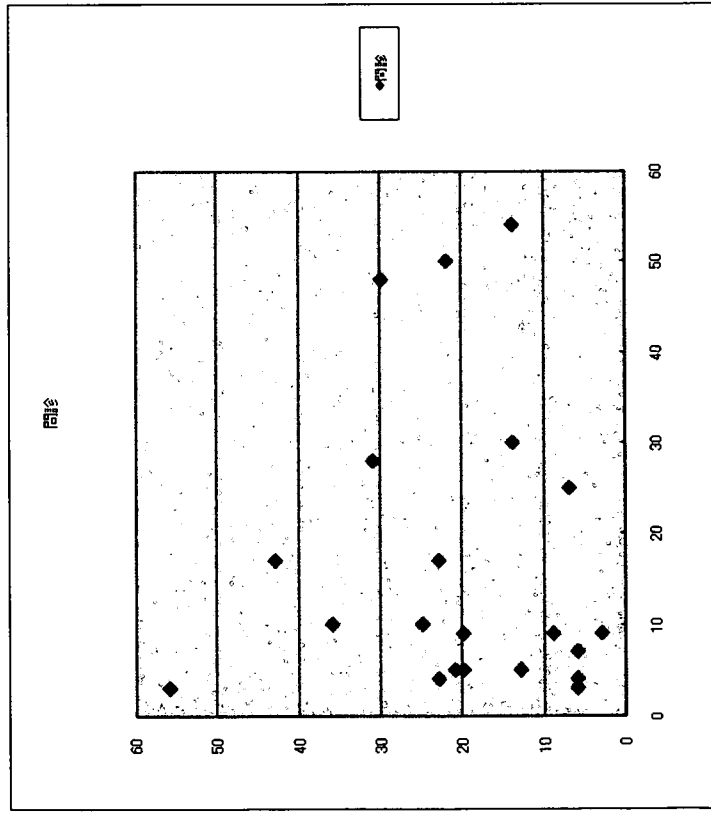


図3 SLE問題の個々の学生の得点と選択数

得点と選択問診数



得点と選択検査数

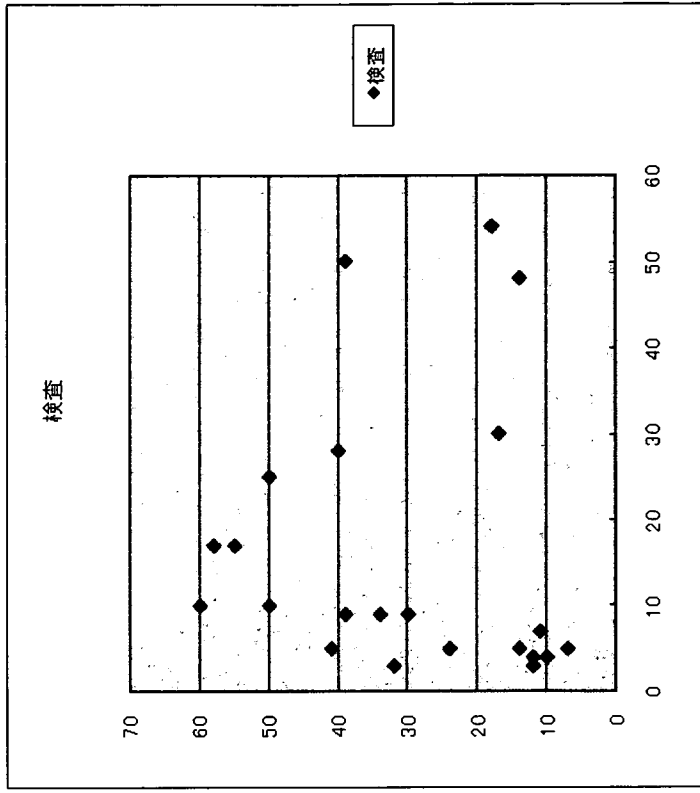
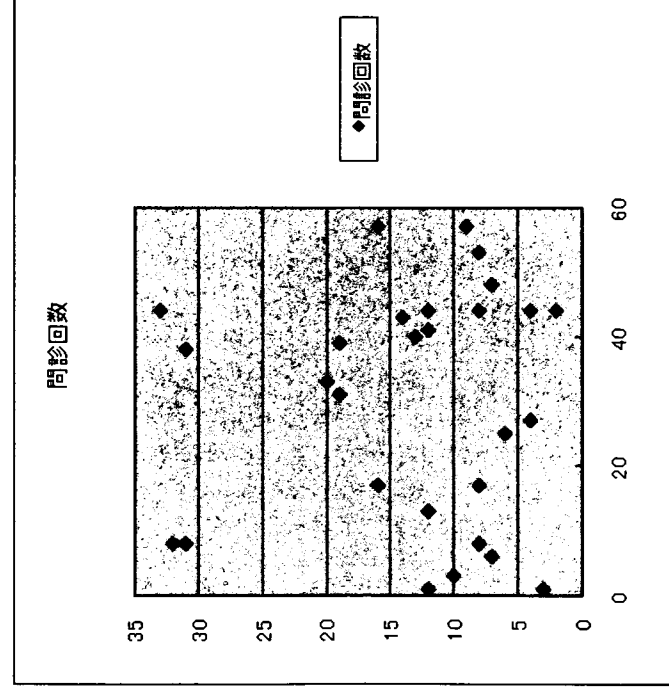


図3 腎盂腎炎問題の個々の学生の得点と選択数

得点と選択問診数



得点と選択検査数

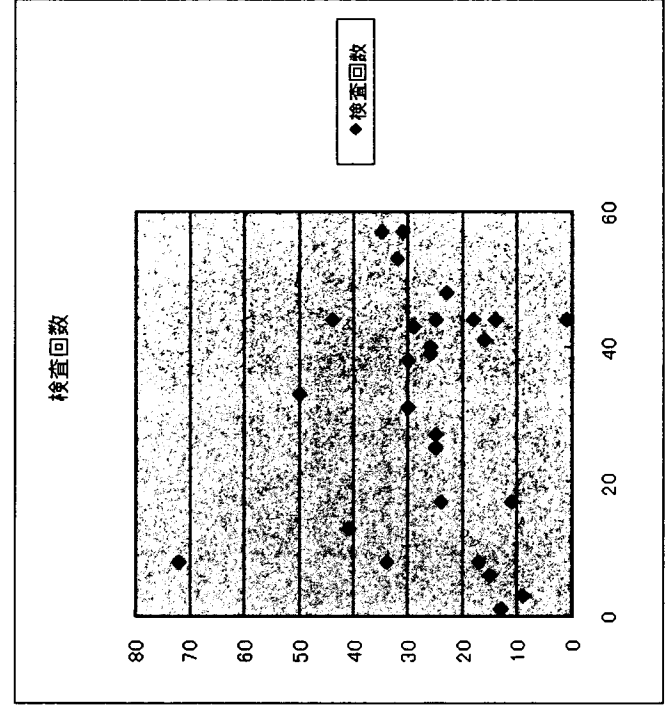


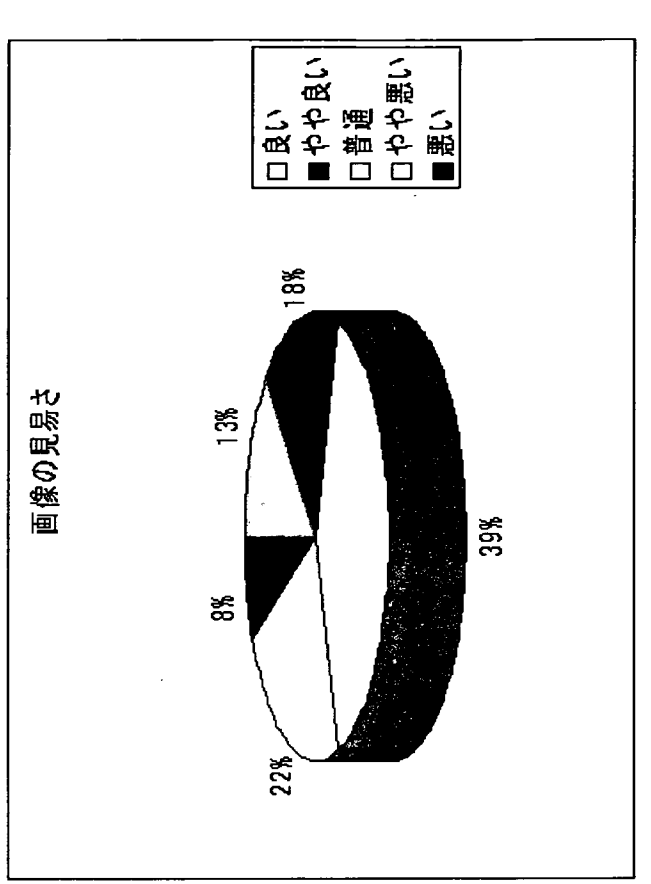
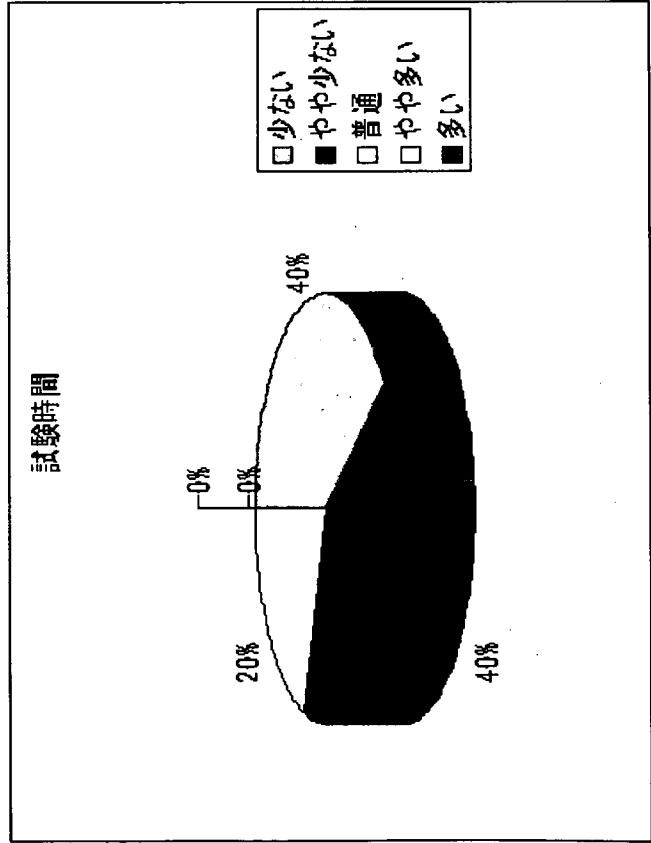
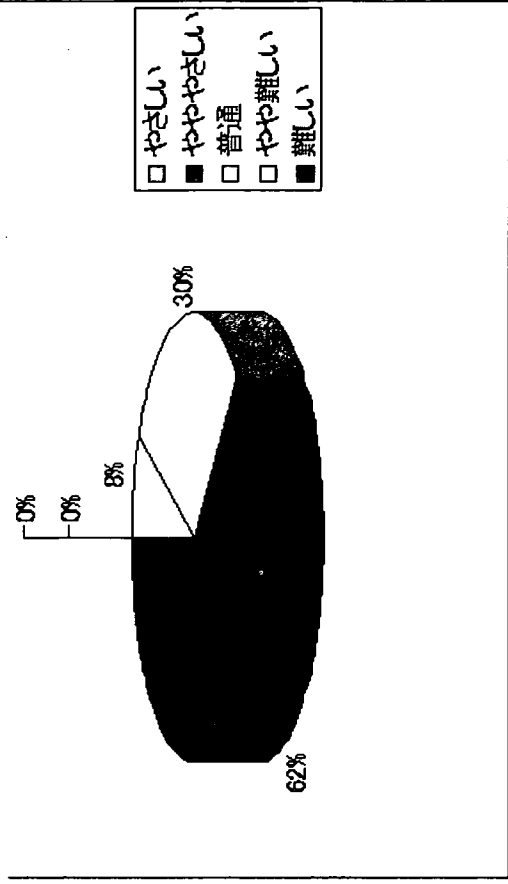
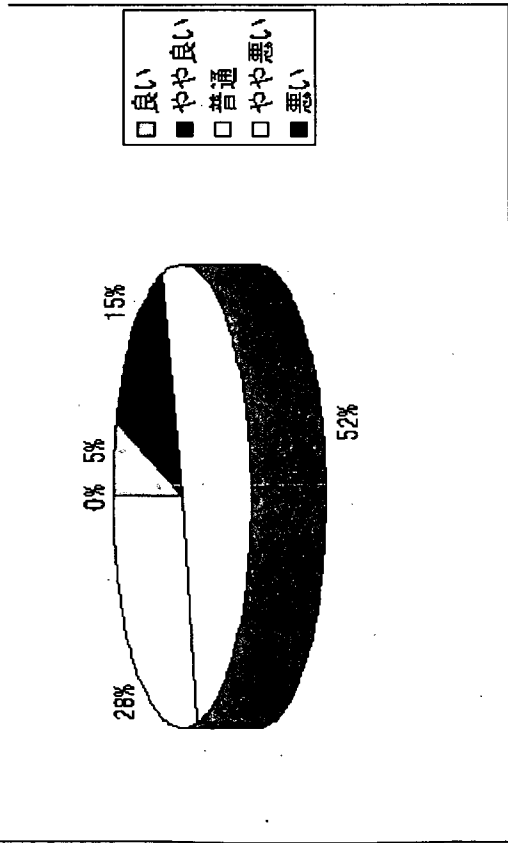
表3 得点と正解項目率の関係

	問診正解率			初めの5問正解率		
	インフルエンザ	SLE	腎盂腎炎	インフルエンザ	SLE	腎盂腎炎
疾患						
学生	0.394±0.192	0.317±0.180	0.424±0.229	0.73±0.254	0.42±0.25	0.49±0.27
医師	0.456±0.158	0.360±0.095	0.486±0.244	0.511±0.176	0.27±0.11	0.5±0.38

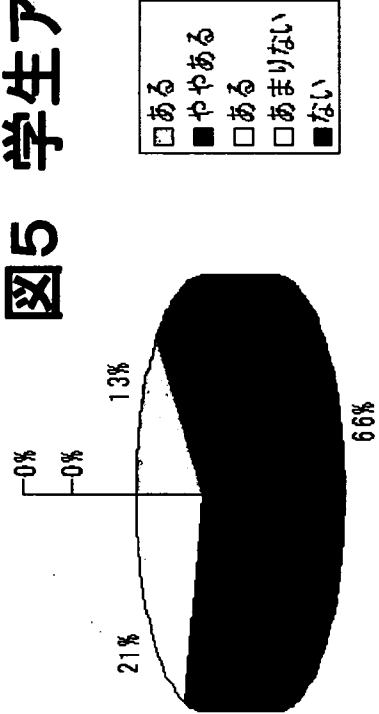
平均±SD

図5 学生アンケートの結果 1

8855の難易度



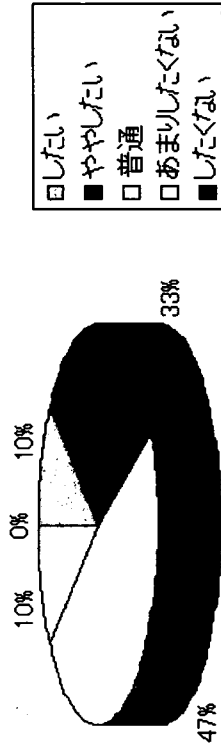
試験への関心度



- ある
- ややある
- ある
- あまりない
- ない

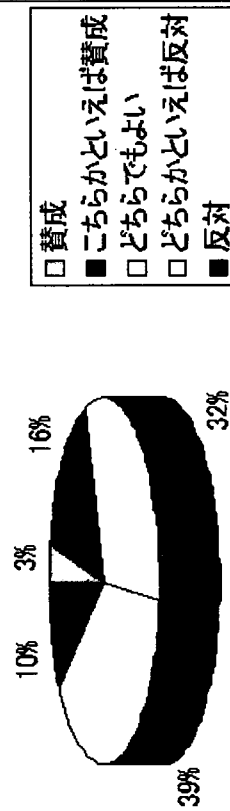
図5 学生アンケートの結果 2

再度この試験を受けたいですか



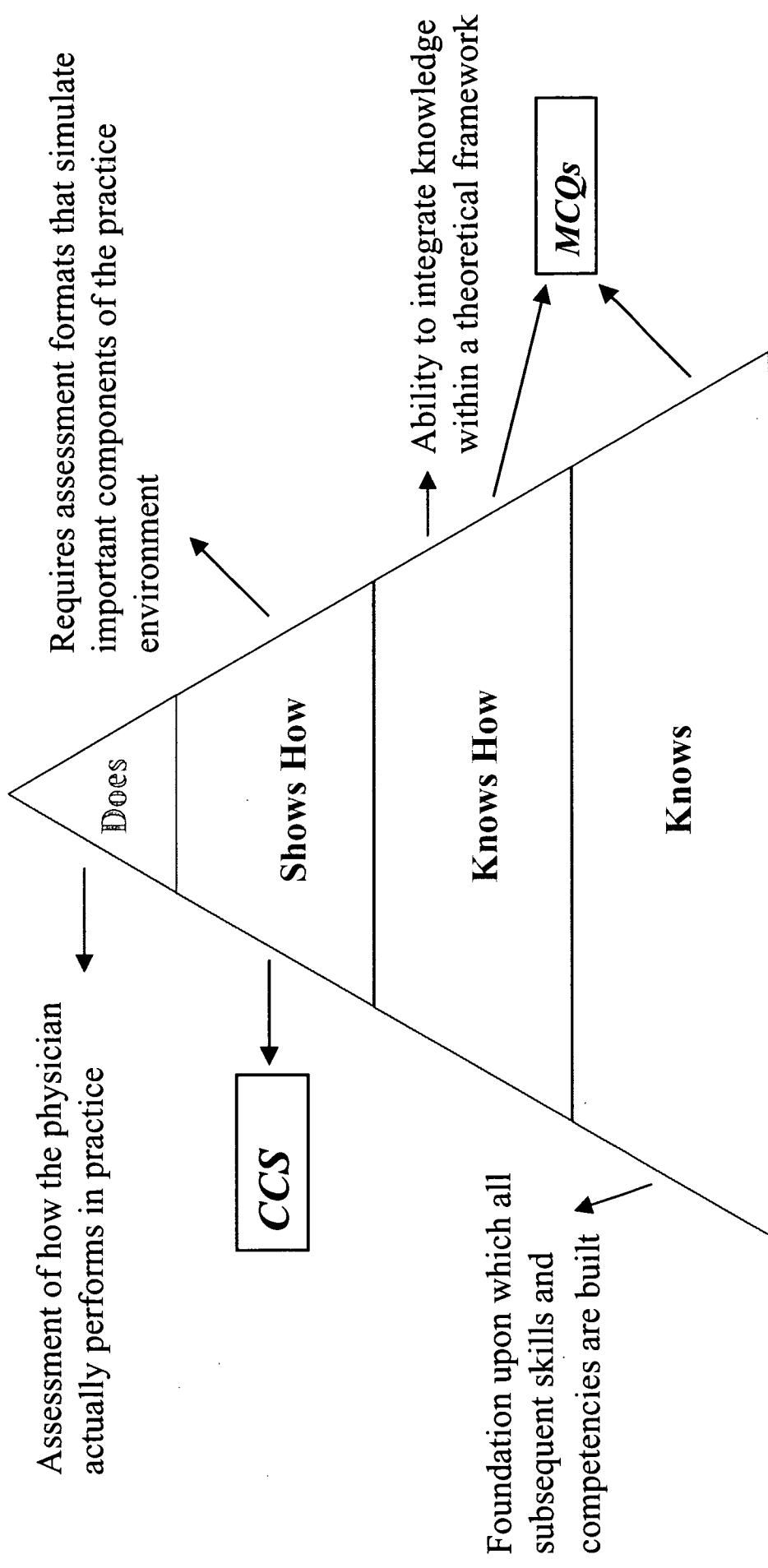
- したい
- ややしたい
- 普通
- あまりしたくない
- したくない

国家試験への導入



- 賛成
- こちらかといえは賛成
- どちらでもよい
- どちらかといえは反対
- 反対

Miller's Taxonomy of Physician Competencies



NBMEの資料からの引用
一部修正

以下NBMEにおける評価方法についての資料

Three Original Scoring Approaches

- Raw score
- Rule-based policy capturing
- Regression-based policy capturing

Regression-Based Policy Capturing

- Experts review and rate hundreds of transaction lists
- Regression equation produced for each case
 - Dependent measure
 - Mean expert rating
 - Independent measures
 - Count of items within each action category
- Algorithms produce scores that approximate the ratings that would have been produced by content experts

Action Categories

- Beneficial Actions
 - Least important
 - More important
 - Most important
- Detractors
 - Non-harmful
 - Risky
 - Extremely Dangerous
- Timing/Sequence

Rating	B_{most}	B_{more}	B_{least}	NH	R	ED	TM
6	4	2	3	2	0	0	3
2	3	1	2	4	1	1	0
7	4	2	3	0	0	0	2
5	3	1	3	2	1	0	1
3	2	2	2	4	1	0	0
9	5	3	3	0	0	0	3
4	2	1	2	3	0	0	1

$$\text{Predicted Rating} = W1*B_{\text{most}} + W2*B_{\text{more}} + \dots + W7*TM$$

Weighted score \approx Predicted Rating

Estimated Regression Weights

Variable	Weight
Beneficial - Most	1.50
Beneficial - More	0.75
Beneficial - Least	0.20
Non-harmful	-0.05
Risky	-1.10
Extremely Dangerous	-2.00
Timing	1.30

$$\text{Weighted score} = 1.5 * B_{\text{most}} + \dots - 2 * ED + 1.3 * TM$$

Example: Rule-based Scoring

Compute Score = 1

IF (B10 > 0) or (B12 > 0)

Score = 3

IF (B3 > 0) and (B4 > 0) and (B5 > 0) and
(B6 > 0) and (B7 > 0) and (B11 > 1) and
((B10 > 2) or (B12 > 2))

Score = 5

IF (B1 > 0) and (B2 > 0) and (B4 > 1) and (B5 > 1) and
(B6 > 0) and (B8 > 1) and (B9 > 1) and (B11 > 1) and
(B12 > 1) and (B13 > 0) and (B14 > 0) and (B20 > 2)

Score = 7

IF (B1 > 0) and (B2 > 0) and (B4 > 1) and (B5 > 1) and
(B6 > 0) and (B7 > 0) and (B8 > 1) and (B9 > 1) and
(B10 > 2) and (B11 > 1) and (B12 > 1) and (B13 > 0)
and (B14 > 0)

Score = 9

B1=History; B2=Physical exam; B3=Gonococcal culture (before antibiotic); B4=Gonococcal culture (time); B5=Gonococcal culture (type); B6=Chlamydia culture; B7=pregnancy test; B8=Blood analysis; B9=Urinalysis; B10=Antibiotic; B11=Antibiotic (time); B12=Gonococcal culture followed by antibiotic; B13=Repeat abdominal exam; B14=Ferrous sulfate

Correlations between ratings of case performance and scores produced with raw, regression-based, and rule-based scoring procedures

Case	Raw Score	Regression-Based Score	Rule-Based Score
1	.76	.81	.77
2	.66	.91	.85
3	.78	.89	.87
4	.80	.88	.84
5	.77	.84	.69
6	.71	.86	.87
7	.54	.79	.79
8	.78	.95	.86

厚生労働科学研究費補助金(医療技術評価総合研究事業)
医師国家試験コンピューター化に関する研究
分担研究報告書

医師国家試験としてのコンピューター試験の適正化

分担研究者

吉岡俊正

東京女子医科大学医学部医学教育学教授

研究要旨

平成 17-18 年度に行った医師国家試験コンピューター化に関する視察結果に基づき、平成 19 年度にコンピューター試験による患者問題解決能力試験システムの至適化を行った。平成 17 年度はドイツ・フランス・英国を視察し、医師国家試験には従来の多肢選択型ではなく画像などを用いた問題解決型試験がコンピューターによる試験では適当であることが明らかになった。平成 18 年度は米国の研修終了後に行う臨床問題解決型コンピューター試験を視察し、前年度の知見を確認するとともに日本では卒前教育のレベルに合わせた問題解決型試験を作成することが国家試験として必要と考えられた。平成 19 年度は問題解決能力試験 (Problem-solving ability test, P-SAT) を開発し評価した。新しい出題フォーマットとして多選択肢問題、ショートエッセイ問題 (短文記入型問題)、語句抽出型問題という 3 つの型を確立した。95 名の学生に行った試験では、多数の問題について専門医と同じ判断をした学生と共用試験 CBT の成績と明らかな相関はなく、問題解決能力評価は通常の CBT とは違う能力 (臨床推論、臨床判断) を評価していることが示唆された。

A. 研究目的

医師に求められる能力として問題解決能力があり、臨床的には臨床推論、臨床判断などで問題発見解決が実践される。医師国家試験では想起的知識だけでなく、問題解決能力の評価が必要とされている。

本研究では医師国家試験コンピューター化にあたり、従来の紙媒体で実施することのできない問題解決能力をコンピューター試験で実施することができるかについて検討した。問題解決能力評価は、Objective structured clinical examination (OSCE) で代表される実技試験などでも実施可能であるが、国家試験としての利便性・再理性・妥当性を考えるとコンピューター試験が適していると考えられる。そこで、問題解決能力評価をどのように行うかについ

での観点に立って、欧州ならびに米国での医師国家試験資格評価について検討し、日本において医学部卒業時に実施するのに適した問題解決能力評価方法を検討した。

B. 研究方法

1) 国外の事例検討

- a. ドイツバイエルン州医師国家試験監督局 (ミュンヘン、平成 17 年 12 月 5 日)
- b. ドイツ国立医師・薬剤師国家試験作政局 (German National Institute fore Assessment of Physicians and Pharmacist. IMPP) (マインツ、平成 17 年 12 月 5 日)
- c. ドイツ厚生労働省 (ボン、平成 17 年