

令和 6 年 6 月 13 日現在

機関番号：82616

研究種目：基盤研究(C)（一般）

研究期間：2021～2023

課題番号：21K02849

研究課題名（和文）非教科型テストによる知識の活用能力の評価に関する研究

研究課題名（英文）Study on evaluation of ability to apply knowledge by non-subject-based tests

研究代表者

椎名 久美子（Shiina, Kumiko）

独立行政法人大学入試センター・研究開発部・教授

研究者番号：20280539

交付決定額（研究期間全体）：（直接経費） 1,100,000円

研究成果の概要（和文）：非教科型テストの各問題項目の正誤データに認知診断モデルを適用して、各問題項目で正答を得るために必要なスキルを統計的手法によって抽出すると共に、解答過程に関する知見を援用してスキルの意味づけを行った。各スキルの習得が各問題項目の正答確率に及ぼす効果を推定すると共に、各受験者のスキルの習得状態を推定した。その結果、同程度のテスト得点でも、スキルの習得状態が異なる受験者が混在していることが示唆された。

研究成果の学術的意義や社会的意義

認知診断モデルを適用する前提で設計されていない非教科型テストであっても、テストの正誤データから受験者のスキル習得状態に関する情報を引き出せる可能性が示されたことは、スキルの習得状態を診断する非教科型テストの今後の作成に生かせる成果である。大学で学ぶためのスキルの習得状態を診断する様々な非教科型テストの開発への応用が期待されるという意味で、円滑な高大接続（高校と大学の教育の接続）に役立つ成果が得られた。

研究成果の概要（英文）：A cognitive diagnostic model was applied to examinees' response data of a non-subject-based test. The skills required to solve each item were specified using statistical methods and the skills were interpreted using insight from research on solving process. The effect of acquiring each skill on the probability that an examinee will answer each item correctly was estimated, and the skill acquisition status of each examinee was estimated. The results suggested that even examinees with similar test scores had different skill acquisition statuses.

研究分野：空間認識力の評価

キーワード：非教科型テスト 認知診断モデル 解答過程

1. 研究開始当初の背景

令和3年度(2021年度)大学入学者選抜では、総合型選抜や学校推薦型選抜において、書類審査や面接、調査書などを材料としつつ、大学教育を受けるために必要な知識・技能、思考力・判断力・表現力等も適切に評価することが求められるようになった(文部科学省, 2020)。総合型選抜や学校推薦型選抜による志願者は高校までの履修教科・科目や学習歴が多様であり、大学で学ぶための思考力や判断力を評価する際に特定の科目の知識を前提とするのが難しい。基礎的な知識を活用して高度な知識を獲得したり、基礎的な知識をもとにデータを解釈したりするような能力を非教科型のテストで評価することが出来れば、入学者選抜にとどまらず、大学入学後のリテラシー教育との接続にも役立つと期待された。このような背景のもとで、多枝選択式の非教科型テストの解答データから、各受験者の知識の習得状態を推定する研究に取り組むに至った。

2. 研究の目的

当初の研究目的は、これまでに開発された非教科型テストの問題の解答過程や構成要素を分析して、知識の活用に重点をおいたテストに変更するための手がかりを得ることであった。そのために、非教科型テストの各問題項目の正誤データから、正答を得るためのスキルの習得状態を推定することを試みた。研究で得られた成果は、どのような特徴の問題がどのような能力の習得を評価するのに適しているのかを検討する材料になることが期待される。

3. 研究の方法

認知診断モデル(cognitive diagnostic models; CDM)は、ある学習領域の習得に必要な要素に関して、テストの解答データから各解答者の習得状態を推定する手法として着目されている(Leighton and Gierl, 2007)。本研究では、CDMのひとつであるG-DINAモデル(generalized-DINA model; de la Torre, 2011)を、非教科型の空間テストである切断面実形視テスト(CEEB, 1939)(Mental Cutting Test, 以下MCTと略記)の解答データに適用した。通常、認知診断モデルは、どの問題項目でどのスキルが必要かを示すQ行列に基づいて設計されたテストに適用されるが、本研究では、MCTの25問に解答した大学生588名の正誤データ(Saito et al., 1998)からQ行列を作成して、G-DINAモデルのパラメータを推定した。そして、正答を得るために必要な要素の習得状態を各受験者について推定して、テスト得点との関係を分析した。一連の分析には、統計解析環境RのGDINAパッケージ(Ma et al., 2020)およびcdmToolsパッケージ(Nájera et al., 2024)を用いた。

4. 研究成果

MCTの正誤データに関して次元構造の分析を行った結果、3つのスキルが抽出され、表1に示すQ行列が示唆された。表1は、3つのスキル(A1, A2, A3)のうち、各問題項目(No. 1~25)で正答を得るために必要なスキルに1が入っている。各問題項目を構成する立体図形や選択枝の特徴や、MCTの解答過程に関する知見(Saito et al., 1996)に基づいて検討した結果、A1は3次元形状の理解、A2は立体と切断面の位置関係の理解、A3は切断形状の部分的および量的な特徴の理解と解釈した。

G-DINAモデルでは、受験者が必要な複数のスキルのうちの一部を習得している場合にその項目で正答する確率が高くなることが表現される。MCTのQ行列(表1)では、25個の問題項目のうち、正答を得るために3つのスキルすべてが必要な問題項目は無い。必要なスキルの数は2または1である。

必要なスキルの数が2の問題項目については、各スキルの習得の有無の組合せ(スキル状態)は4パターンとなる。パターン l ($l = 1, 2, 3, 4$)は、問題項目 j で正答するために必要な1番目のスキルの習得の有無を示す α_{j1} (1または0)と2番目のスキルの習得の有無を示す α_{j2} (1または0)の4通りの組合せに対応し、パターン l のスキル状態の受験者が問題項目 j で正答を得る確率 P_{lj} は(1)式で表現される。

$$P_{lj} = \delta_{j0} + \delta_{j1}\alpha_{l1} + \delta_{j2}\alpha_{l2} + \delta_{j12}\alpha_{l1}\alpha_{l2} \quad (1)$$

必要なスキルの数が1の問題項目については、スキル状態は2パターンとなる。パターン l ($l = 1, 2$)は、問題項目 j で正答するために必要なスキルの習得の有無を示す α_{j1} (1または0)で表現

表 1: 切断面実形視テストの Q 行列

Item No.	A1	A2	A3
1	1	0	1
2	1	1	0
3	0	1	0
4	1	0	0
5	0	1	0
6	1	1	0
7	1	1	0
8	0	1	0
9	0	1	0
10	1	1	0
11	1	0	1
12	0	1	1
13	1	0	1
14	0	1	1
15	0	1	0
16	1	0	1
17	1	0	1
18	1	0	1
19	1	0	1
20	1	0	1
21	0	0	1
22	0	0	1
23	0	0	1
24	0	0	1
25	1	0	1

され、パターン l のスキル状態の受験者が問題項目 j で正答を得る確率 P_{lj} は(2)式で表現される。

$$P_{lj} = \delta_{j0} + \delta_{j1}\alpha_{l1} \quad (2)$$

(1)式と(2)式における δ_{j1} および δ_{j2} は、各スキルの習得が正答を得る確率に及ぼす主効果、(1)式における δ_{j12} は、必要な2つのスキルの習得が及ぼす交互作用を意味する。(1)式と(2)式における δ_{j0} は、必要なスキルをどれも習得していない場合に正答を得る確率を意味する。

MCTの正誤データに表1のQ行列を用いたG-DINAモデルを適用して、問題項目25問の項目パラメータ((1)および(2)式の $\delta_{j0}, \delta_{j1}, \delta_{j2}, \delta_{j12}$)を推定した結果が表2である。

図1に、MCTの問題項目17に関して、必要なスキルA1(3次元形状の理解)とA3(切断形状の部分的および量的な特徴の理解)の習得の4つのパターン(00, 10, 01, 11:左の数字はA1の習得の有無,右の数字はA3の習得の有無を示す)の受験者が正答を得る確率を、表2の項目パラメータを用いて算出した結果を示す。スキルを1つしか習得していないパターン(10および01)よりも、2つとも習得しているパターン(11)のほうが、正答を得る確率が高いことが表現されている。

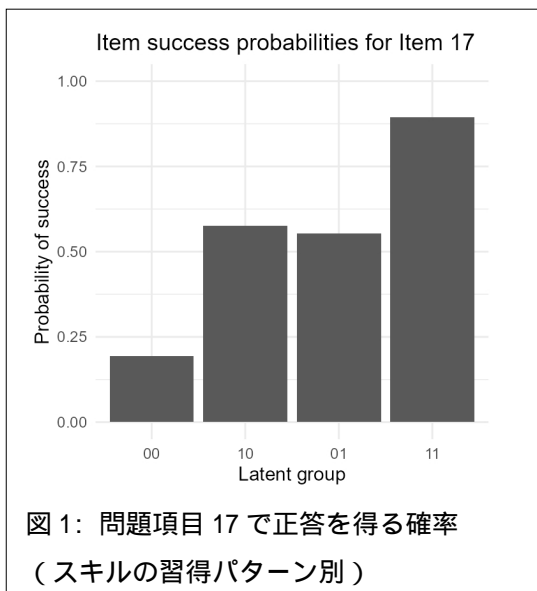


表2: 各問題項目の項目パラメータの推定値

Item	δ_{j0}	δ_{j1}	δ_{j2}	δ_{j12}
1	0.771	0.191	0.193	-0.191
2	0.344	0.232	0.263	0.053
3	0.939	0.061		
4	0.452	0.432		
5	0.706	0.239		
6	0.550	0.264	0.411	-0.264
7	0.321	0.287	0.263	0.040
8	0.755	0.223		
9	0.210	0.311		
10	0.421	0.253	0.233	0.006
11	0.675	0.271	0.205	-0.186
12	0.446	0.307	0.120	-0.017
13	0.214	0.103	0.126	0.144
14	0.236	0.241	0.197	0.092
15	0.513	0.418		
16	0.442	0.129	0.006	0.341
17	0.194	0.383	0.360	-0.043
18	0.256	0.426	0.454	-0.255
19	0.296	0.248	0.354	-0.071
20	0.423	0.294	0.381	-0.128
21	0.439	0.377		
22	0.316	0.512		
23	0.226	0.286		
24	0.311	0.300		
25	0.088	0.106	0.207	0.288

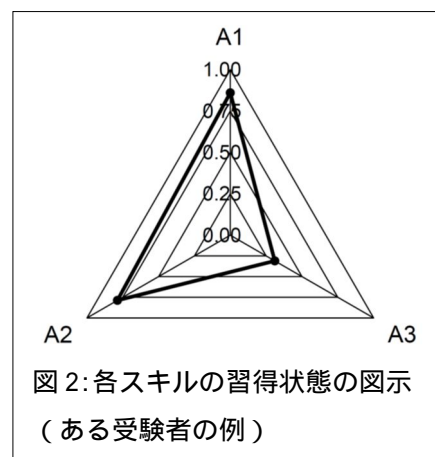
MCT25問のうち、問題項目17のように、多くのスキルを習得するほど正答を得る確率が高くなる問題項目が多くみられる一方で、スキルの習得が正答を得る確率に及ぼす効果が少ない問題項目もある。これは、表1のQ行列に改善の余地があることを示唆しており、今後、複数の専門家との再検討を要する課題である。

各受験者がA1~A3の各スキルを習得している確率を、GDINAパッケージを用いて推定した。図2は、ある受験者についてレーダーチャートを用いて各スキル習得確率を示した例である。この受験者は、A1(3次元形状の理解)とA2(立体と切断面の位置関係の理解)は習得できている可能性が高いが、A3(切断形状の部分的および量的な特徴の理解)を習得できている可能性は低いことが示されている。

各スキルの習得確率が0.5を超えていれば1、そうでなければ0として習得状態を2値化すると、A1~A3の3つのスキルの習得状態の組合せは $2^3=8$ パターンありうる。表3は、MCTの各得点範囲におけるスキルの習得パターン(attribute patterns)の人数を示すものである。表3の習得パターンは、各スキルの2値化された習得状態を左から3つ並べて表記している。

表3には、高得点範囲になるほど、多くのスキルを習得した受験者が増えることが示されている。また、同じ得点範囲内に、正答を得るために必要なスキルの習得状態が異なる受験者が存在しており、得点だけでは得られない情報として、スキルの習得状態に関する情報が得られたことが示唆される。

本研究では、認知診断モデルを適用する前提で設計されていない非教科型テストに関して、Q



行列の推定を行って G-DINA モデルを適用した。Q 行列には改良の余地があるものの、受験者のスキル習得状態に関する情報を引き出すことにある程度成功したと言える。これは、非教科型テストを診断的に使うことに繋がる成果である。今後の展望として、スキルの習得が正答確率に大きく寄与する問題項目に着目して問題の特徴や構成要素を類型化することで、スキルの習得状態を診断できる非教科型テストの問題の安定的な作成に繋がる事が期待される。

表 3: 切断面実形視テストの各得点範囲におけるスキル習得パターン(attribute patterns)の人数

MCT score ranges	Latent groups (attribute patterns)							Total
	000	100	010	001	110	011	111	
0-10	67	5	0	0	0	0	0	72
11-13	50	29	4	6	1	0	0	90
14-15	8	42	3	19	13	4	0	89
16-17	0	28	1	7	27	19	3	85
18-19	0	2	0	0	19	15	47	83
20-21	0	0	0	0	1	3	73	77
22-25	0	0	0	0	0	0	92	92
Total	125	106	8	32	61	41	215	588
Proportion(%)	21.3	18.0	1.4	5.4	10.4	7.0	36.6	

<引用文献>

文部科学省: 令和3年度大学入学者選抜実施要項(2020年6月19日). https://www.mext.go.jp/content/20200624-mxt_daigakuc02-100001251_1.pdf (最終閲覧2024年6月13日).

Leighton, J., Gierl, M. (Eds.): *Cognitive diagnostic assessment for education: Theory and applications*. New York, NY: Cambridge University Press (2007).

de la Torre, J.: The generalized DINA model framework. *Psychometrika* **76**, 179-199 (2011). doi:10.1007/s11336-011-9207-7

CEEB Special aptitude test in spatial relations, developed by the College Entrance Examination Board, USA (1939).

Saito, T., Suzuki, K., Jingu, T.: Relations between spatial ability evaluated by a mental cutting test and engineering graphics education. *Proceedings of the 8th International Conference on Engineering Computer Graphics and Descriptive Geometry*, 231-235 (1998).

Ma, W., de la Torre, J.: GDINA: An R package for cognitive diagnosis modeling. *Journal of Statistical Software* **93**(14), 1-26 (2020). doi:10.18637/jss.v093.i14

Nájera, P., Sorrel, M. A., Abad, F. J.: cdmTools: Useful tools for cognitive diagnosis modeling. R package version 1.0.5.(2024). <https://cran.r-project.org/web/packages/cdmTools/> (accessed on June 13, 2024)

Saito, T., Shiina, K., Suzuki, K. Jingu, T.: Spatial ability evaluated by a mental cutting test. *Proceedings of the 7th International Conference on Engineering Computer Graphics and Descriptive Geometry*, 569-573 (1996).

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計1件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 Kumiko Shiina
2. 発表標題 Applying a Cognitive Diagnosis Model to Mental Cutting Test
3. 学会等名 The 21st International Conference on Geometry and Graphics (ICGG 2024) (国際学会)
4. 発表年 2024年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------