

クリティカルシンキングの適応型テストのアイテムバンク構築

Construction item bank for critical thinking adaptive test development

若山 昇^{*1}
Noboru WAKAYAMA^{*1}
^{*1} 帝京大学/CRET

宮澤 芳光^{*2}
Yoshimitsu MIYAZAWA^{*2}
^{*2} 東京学芸大学

梶谷 真司^{*3}
Shinji KAJITANI^{*3}
^{*3} 東京大学

植野 真臣^{*4}
Maomi UENO^{*4}
^{*4} 電気通信大学

^{*1}Teikyo University/CRET ^{*2}Tokyo Gakugei University
^{*3}The University of Tokyo ^{*4}The University of Electro-Communications
Email: Wakayama.class@pobox.com

あらまし：現代は様々な情報があふれており、十分に吟味した情報を体系的に理解し、的確な意思決定を行うためには、クリティカルシンキング(CT)は不可欠である。本研究の目的は、CTの適応型テストの開発のために、アイテムバンクの構築を試みるものである。その結果、CTの3尺度（分析、推論、読解）は、相関、散布図、情報量などを分析したところ、概ね独立であり妥当であると考えられた。

キーワード：クリティカルシンキング、適応型テスト、項目反応理論、試験、評価

1. はじめに

1.1 クリティカルシンキングの重要性

現代は様々な情報があふれており、十分に吟味した情報を体系的に理解し、的確な意思決定を行うためには、クリティカルシンキング(CT)は不可欠である。社会人基礎力(経済産業省 2006)や学力(文部科学省 2008)、さらには 21 世紀型スキル(Griffin 他 2011)においても、その重要性が謳われている。なお、クリティカルシンキング(CT)とは、先入観に囚われず、論理的に考え、合理的な決定を導き出す能力と意思である⁽¹⁾。

1.2 クリティカルシンキング能力の評価

CTの能力評価については、海外のテストを利用したり(平山他 2011)、国内でも能力評価の研究は存在する(楠見他 2010)。しかし、CTの試験では読解時間に加えて、考える時間が十分に必要であるため時間がかかってしまうのが現状である。

1.3 項目反応理論について

項目反応理論(IRT)では受検者ごとに瞬時に最適な問題を提示可能となり、測定精度を落とさず、時間を短縮できる。一方、既存のテストでは、同じテストを受検者した同士でしか能力を比較することができないが、IRTでは、異なる項目で構成されたテストでも同一の尺度上で受検者を評価できる。

1.4 研究の意義

CTの重要性の高まりとともに、学生・社会人のCT能力を測定する必要性が高まっている。短時間で効率的に能力を測定するためにCTの適応型テストの技術開発が必要となる。このため、アイテムバンクの構築する不可欠となる。

1.5 研究目的

本研究の目的は、CTの適応型テストの開発のために、アイテムバンクの構築を試みるものである。

2. 研究方法

東京及び近郊の大学生 736 人を対象とした。18 歳から概ね 22 歳で、解答に不備がある者と留学生を除外している。成績に無関係で、教育・研究目的以外には使用せず個人名は特定されないことを説明した。

CTの能力評価には、行動観察や口頭試問などもあるが、本研究では多肢選択型試験で測定できるものと仮定しペーパー試験を実施した。設問は、CTの定義にできるだけ沿って、独自に作成したもの及び公務員試験など既存の問題の中からCTに必要な①分析的思考力、②論理・推論能力、③読解・理解能力を測る3尺度で構成した。作問の際は、大学教員かつ博士号を持ち教育に関する研究を行う4人の研究者が議論を重ね、問題を策定した。

各尺度毎に、5問を共通問題とし15問の試験を5組作成し、各組毎に50人程度から解答を得た。なお、①分析と③読解の被検者は同一の107人であり②の共通問題を含めて終日試験に解答してもらった。②推論の632人は、授業で試験を1組ずつ行って集計した。データをIRTPRO3及びSPSS20.0で解析した。

3. 結果

3.1 尺度の相関と困難度・識別力

個人の能力における各尺度間の相関と散布図は以下のとおり(表1, 図1~3)。また、各尺度毎の困難度・識別力の散布図は以下のとおり(図4~6)。

	Θ1	Θ2	Θ3
Θ1			
Θ2		.568**	.570**
Θ3			.423**

**: $p < .01$

3.2 各尺度のテスト情報量と標準誤差

各尺度のテスト情報量(I)と標準誤差(S. E.)は以下のとおり(表 2, 図 7~9)

表2 尺度のテスト情報量と標準誤差

	情報量(max)	標準誤差	困難度
①分析的思考力	11.65	0.29	-0.20
②論理・推論能力	17.44	0.24	0.20
③読解・理解能力	10.15	0.31	-0.70

4. 考察

各尺度(①分析的思考力, ②論理・推論能力, ③読解・理解能力)の相関は大きくなく, 散布図から2元性は確認されず, 偏りも大きくはない. 困難度・識別力の散布図は, その性質上の要因もあるが困難度0付近を中心に正規分布のような形状が確認された. また, 各尺度のテスト情報量(I)と標準誤差(S. E.)については, 被検者数が十分に大きいとはいえないために, グラフの形状及びその値には限界があるものの概ね妥当な範囲におさまっていた. また,

②推論では被検者が他の尺度より多いため, グラフには若干の歪みが見られるが誤差は小さめ(0.24)となった. これらにより, 各尺度は概ね独立的であり妥当であることが示唆された.

5. おわりに

今後は被検者を増やし, 識別力の極端に低いものを排除し検証することなどにより, さらに的確なアイテムバンクを構成する必要があるだろう. 本研究は科研費(A)15H01772, (C)15K01088 及び CRET(教育テスト研究センター)の助成を受けている.

参考文献

- (1) 若山昇: “大学におけるクリティカルシンキング演習授業の効果 -クリティカルシンキングに対する志向性と認知欲求の変化から-”, 大学教育学会, Vol.31, No.1, pp.145-153 (2009)

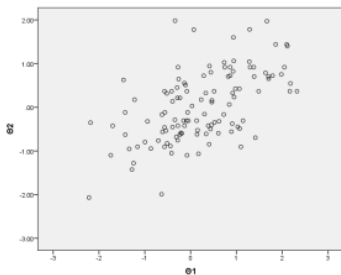


図1 ①分析的思考力と②論理・推論能力

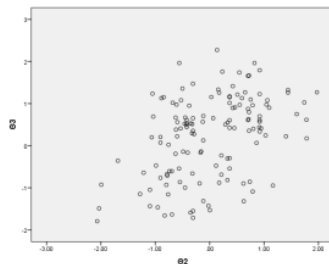


図2 ②論理・推論能力と③読解・理解能力

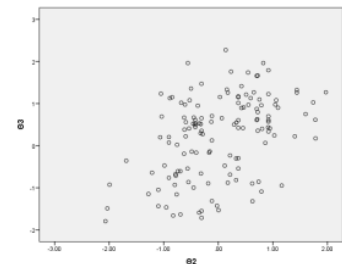


図3 ③読解・理解能力と①分析的思考力

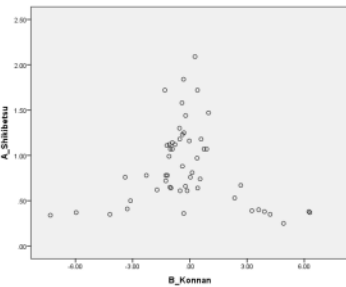


図4 ①分析的思考力の困難度と識別力

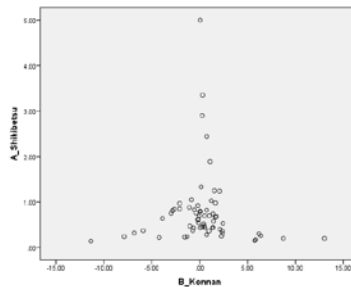


図5 ②論理・推論能力の困難度と識別力

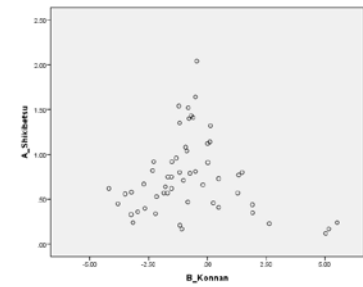


図6 ③読解・理解能力の困難度と識別力

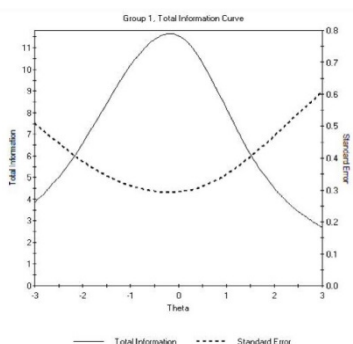


図7 ①分析的思考力の情報量と標準誤差

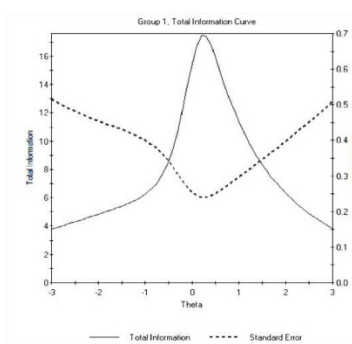


図8 ②論理・推論能力の情報量と標準誤差

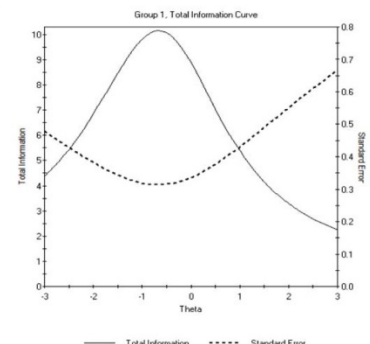


図9 ③読解・理解能力の情報量と標準誤差