



## 認知的熟慮性検査（CRT）における項目間等価性および呈示順序・教示効果の検証：大学生集団実験による検討

著者	原田 佑規, 原田 悦子, 須藤 智
雑誌名	筑波大学心理学研究
巻	56
ページ	27-34
発行年	2018-08-31
URL	<a href="http://hdl.handle.net/2241/00153637">http://hdl.handle.net/2241/00153637</a>

# 認知的熟慮性検査 (CRT) における項目間等価性 および呈示順序・教示効果の検証 ——大学生集団実験による検討——<sup>1)</sup>

筑波大学人間系 原田 佑規・原田 悦子

静岡大学大学教育センター 須藤 智

Examination of Cognitive Reflection Test (CRT): Homogeneity of questions, order effects, and effects of instruction

Yuki Harada and Etsuko T. Harada (*Faculty of Human Science, University of Tsukuba, Tsukuba 305-8572, Japan*)

Satoru Suto (*Higher Education Development Center, Shizuoka University, Shizuoka 422-8529, Japan*)

The Cognitive Reflection Test (CRT: Frederick, 2005) is a test designed to measure individual differences in intuitive-reflective thinking styles, and consists of three questions in a fixed order. Although Toplak, West, & Stanovich (2014) have proposed four additional questions, their validity as substitutes was not fully examined. This study investigated (1) the equivalences of the original and new questions, (2) order effects for the three original questions, and (3) the dependency of the CRT across instruction conditions (i.e., standard, under time pressures or stressing accuracy). More specifically, we administered a paper-and-pencil version of the CRT to 128 undergraduate students and analyzed both task performances and confirmation ratings. The results indicated (a) that three of the new questions are not suitable substitutes for the original CRT questions, (b) that order effects in terms of independent item-specific characteristics were observed only for the confirmation ratings, and (c) that the instructions had no effect on either performance or confirmation ratings. Finally, we discussed the importance of investigating both the CRT's methodology and the cognitive processes underlying the CRT.

**Key words:** Cognitive Reflection Test, confirmation rating score, dual process theory

Kahneman (2003) は、人の思考には、迅速かつ自動的で、ヒューリスティックな処理を行うシステム 1 と、時間は要するが意識的で分析的なシステム

2 があるとする二重過程理論を提案した。こうした二重過程は記憶研究でも、親近性 (familiarity) に依拠する処理の流暢性と意識的顕在的な処理に基づく記憶想起の二重過程が提案されている (Jacoby, Kelley, & McElree, 1999)。特に、問題解決の文脈において、システム 1 は、直感的で時にバイアスのかかった判断を導き、システム 2 は、熟慮的で正確な判断を導くと考えられ (Stanovich & West, 2000)、二重過程理論は、われわれの非合理的判断

1) 本研究は、科研費 (16H02053) の助成を受けて行われた。本研究の実施にあたり、筑波大学大学院人間総合科学研究科の田中伸之輔氏に多大なご協力をいただきました。この場を借りて、御礼申し上げます。  
連絡先: ☒ etharada@human.tsukuba.ac.jp (原田悦子)

を説明する有力な仮説の一つとなっている。

こうした直感的・熟慮的な判断は思考傾向の個人差を測定するためにも用いられている（たとえば認知バイアス課題：Kahneman & Tversky, 1982）。Frederick（2005）は極めて簡易な指標として認知的熟慮性課題（Cognitive Reflection Test: CRT）を提案した。CRTは特定の文脈の文に埋め込まれた計算問題3問から構成されており、それぞれの問題には、直感的な回答（典型的誤答）と熟慮的な回答（正答）が設定されている（Table 1, 問1から3）。例えば問1「バットとボールは、合わせて1100円です。バットはボールよりも1000円高いです。では、ボールはいくらでしょう」という問題に対し、直感的には二つの数の差分「100円」を回答しがちであるが、数学の問題として考えれば「50円」という熟慮的な回答が求められる。これら3問の合計値としてのCRT得点は、認知バイアス課題や3段論法課題の得点と正の相関関係にあり（ $r_s = .42, .36$ ; Toplak, West, & Stanovich, 2011）、CRTは認知的熟慮性を簡便に計る指標として広く受け入れられてきている。

こうした状況に対し、CRTの問題数が3問のみであり、測定の反復に適さないことから、等価性のある問題追加の必要性が主張されるようになり、Toplak, West, & Stanovich（2014）は新たにCRT4問を提案した（Table 1, 問4から7）。しかしこれら追加4問が、オリジナルの3問と同じ、認知的熟慮性を測定しているか否かについては、これまで提案者自身による相関分析以外に検証がなされていない。またオリジナルのCRT3問については、先行研究において邦訳とその実施結果が報告されているが（松原他, 2015）、追加4問については、いまだ日本語を用い、日本人参加者による測定結果についてデータが存在しない。そこで、本研究の第1の目的は、日本語・日本人参加者を対象として、CRT7問を実施し、オリジナル3問と追加4問の等価性を検討することであった。

また思考傾向・問題解決過程の尺度としての妥当性の観点から、オリジナル3問についても等価性、あるいはそれらが何を測定しているのかについての検証についても十分には行われていないと考えられる。CRTのオリジナル3問の結果は、1問1点と

Table 1  
CRTの問1から7までの詳細

番号	問題の詳細
1	バットとボールは、合わせて1100円です。バットはボールよりも1000円高いです。では、ボールはいくらでしょう。 正答：50円 典型的誤答：100円
2	5台の機械は、5分間で5個のおもちゃを作ります。では、100台の機械が100個のおもちゃを作るのには、何分かかりますか？ 正答：5分 典型的誤答：100分
3	池にスイレンの葉が浮かんでいます。葉の面積は毎日倍になります。葉が池を覆いつくすのに48日かかりました。では、半分を覆うまでには何日かかったでしょう？ 正答：47日 典型的誤答：24日
4	太郎は6日間で1タルの水を飲みます。花子は12日間で1タルの水を飲みます。では、太郎と花子が二人で同じタルから水を飲む場合、空になるまでに何日かかるでしょう？ 正答：4日 典型的誤答：9日
5	次郎の成績は、クラスで上から15番目、下から15番目でした。このクラスの人数は何人でしょう？ 正答：29人 典型的誤答：30人
6	ある人は、6000円で買った子豚を7000円で売りました。そして、その子豚を8000円で買い戻し、最終的に9000円で売りました。その人はいくら儲けたでしょう？ 正答：2000円 典型的誤答：1000円
7	2008年の初め、三郎は80万円を株に投資することにしました。6か月後の7月17日に、株は50%下がりました。しかし幸運にも7月17日から10月17日にかけて彼が買った株は75%上がりました。この時点で、正しいのは次のうちどれでしょう？ (a) 三郎は損も得もなかった (b) 三郎は得をした (c) 三郎は損をした 正答：c 典型的誤答：b

Note. 問1から3は、Frederick（2005）のオリジナル問題であり、問4から7は、Toplak et al.（2014）の追加問題に基づいている。原典は英語かつ欧米社会的な用語が用いられていたため、問題は、筆者らによって、日本語に翻訳され、かつ日本文化に合わせて変換された。

して採点・合計される（Frederick, 2005; Simonovic, Stupple, Gale, & Sheffield, 2017; Toplak et al., 2011）が、この処理には、すべての問題の等価性が仮定されている。しかしながら、いくつかの先行研究においては、問1の正答率は問2と3と比べて低かったことが報告されている（Campitelli & Gerrans, 2014; Travers, Rolison, & Feeney, 2016）。この原因として、問1に固有の特性があり、問1と問2、3が異なる思考の属性を測定している可能性も考えられる。

この結果については、問の呈示順序が影響を与えている可能性も否定できない。CRTを利用したほとんどの研究では、問題の呈示順序が固定されている可能性が高い。実際、CRTを使用した多くの研究では、問題の呈示順序に関する記述がなく（e.g., Cueva et al., 2016; Pennycook, Cheyne, Koehler, & Fugelsang, 2015）、多くの研究で当然のようにオリジナルの呈示順序のまま実施されている可能性が極めて高い。しかしながら、CRTの回答者にとって、初めて経験する第1問については、出題意図あるいは文脈を把握しにくく、結果的に直感的な判断が生じやすい可能性が考えられる。あるいは逆に、第2, 第3問目については、先に解いた問題での処理経験の蓄積から、出題意図の推測も一定程度可能になり、その結果として熟慮的な判断が生じやすい可能性も考えられる。

こうした仮説に基づけば、本来3つの問題は等価であるが、問1は最初に出题されるために正答率が低下したと考える可能性を否定できない。一方で、もしそうした呈示順序の効果がまったくないのであれば、問題の間の等価性についてさらに検討をしていくことが必要であろう。したがって、CRTの妥当性として、問題の間の等価性を検討する意味において、問題への回答の差に呈示順序が影響を与えているか否かを検討しておくことは重要である。そこで、本研究では第2の目的として、オリジナルCRTの3つの問題における呈示順序の効果を検証した。

CRTの得点は、ストレス操作などによる状況的要因の影響を受けることが報告されている。例えば、スピーチ課題によるストレス操作はCRTの得点を低下させた（Simonovic et al., 2017）。こうした状況的要因の一つとして、時間圧（time pressure）の影響も考えられる。本来、思考モデルとしての二重過程モデルでは、個々のシステム内におけるスピードと正確さのトレードオフ（speed-accuracy tradeoff）については言及されていないが（Kahneman, 2003）、テスト実施の際に外的条件と

して、時間圧がかけられることにより、トレードオフによって正答率が下がっているなどの現象が現れるならば、CRTが思考スタイルではなく、時間圧に対する耐性の測定をしている可能性も考えられよう。実際、CRTが「何を測定しているのか」についての深い議論はなされていないだけに、状況・環境的要因がCRTに及ぼす影響について明らかにしていくことは、CRTの結果を個人の認知的熟慮性に帰属していくことの限界と可能性を検討していくために必要であろう。そこで、本研究の第3の目的は、教示の操作により、時間圧をかける、あるいは正確さを追求するよう求めることで、CRT得点が変化するか否かを検証することとした。

加えて、これまでのCRTで用いられているのは回答の正否のみであるが、その回答に対して確信度が異なるならば、その背景となっている認知的過程が異なることも考えられる。そこで本研究では、CRTへの回答を求めた後に、その解への確信度についても測定をした（0から10までの11段階評定尺度）。確信度は、問題解決におけるメタ認知の指標としても使用されており（佐藤他, 2017）、確信度の高さは、算数文章題（岡本, 1991）、一般知識問題（篠塚, 1993）等の問題解決の正否と関係をしていることが報告されている。回答の属性と同時にそこへの確信度も合わせて検討することにより、CRTが測定している認知的過程について、より考察を深めることが可能になると考えられる。

以上から、本研究ではCRTの各問題への回答ならびに回答への確信度の両面から、(a) CRTオリジナル3問と追加4問の等価性、(b) オリジナル3問の呈示順序の効果、そして(c) 時間圧および正確さ追求の教示の効果、の三者を検証し、こうしたCRTを用いた研究の基礎データに基づいて、「CRTで何が測られているのか」に関する知見を得ることを目的として実験を行った。

## 方 法

**参加者と実施方法：**大学生128名（男性79名、女性47名、未記入2名；平均年齢19.02歳、SD 1.23）。

**刺激と材料：**調査用紙は、冊子体として配布された。冊子体は表紙に調査の説明（目的、取得されたデータの取り扱い、研究倫理に対する配慮の項目を含む）を記載した。2ページ目には教示、3ページ目にCRTのオリジナル3問、4ページ目に追加4問（全員が同一順）、そして5ページ目にデモグラフィック変数シートがあった。オリジナル3問題の呈示順序効果を検討するために3ページ目について

は6種が作成された。また教示効果を検討するために、通常教示条件ではCRTの回答と確信度の評定の仕方に関する教示が記載されたが、時間圧/正確教示条件では速度/正確性を求める教示が追加され、オリジナルの問題順で提示された (Table 2)。合計8種類の冊子はランダムに配布・実施された。

教示では、CRTの各問題に対して回答すること、その回答に対する確信度を11段階で評定することが求められた。加えて、正確条件教示として、「できるだけ正確にクイズに回答してください」、時間圧条件教示として「できるだけ早くクイズに回答してください」という教示文を追加した。デモグラフィック変数として、回答者の性別、年齢、所属、学年、CRTを見たことがあるか否かの回答を求めた。

**手続き:** 本調査は、大学における一般教養課程科目において、認知心理学における問題解決研究の事例であることを説明した上で、大学の授業終了後に約15分間かけて行われた。講義後に本調査への参加は自由意思による同意に基づくこと、不参加による不利益はないこと、参加後の撤回が可能であること、そして個人は特定されず個人情報保護されることを書面及び口頭で説明した。そして、質問紙への回答をもって調査への協力に同意したものとすることが説明された。調査用紙の回収は、回答が終わった参加者から順に行われた。

## 結 果

各群のデモグラフィック変数の記述統計を Table 2に示す。なおCRTの回答における欠損値は、誤答として処理された。確信度の欠損値は、ペアワイズで分析から除外された。

## オリジナル問題と追加問題の等価性の分析

通常教示条件における問1から7について、正答、典型的誤答、非典型的誤答の比率を求めた (Figure 1)。これらの回答の産出確率について問題の間で比較するため、参加者内計画により、回答種類ごとに問題(7)を要因とした1要因の分散分析を実施したところ、正答の確率については問題の効果が有意であり [ $F(6, 86) = 4.73, p < .0001, \eta_p^2 = 0.05$ ]、多重比較の結果、問4・問5は問2・問7よりも有意に正答率が高かった。典型的誤答の確率については、問題の間の効果が有意であり [ $F(6, 86) = 7.00, p < .0001, \eta_p^2 = 0.08$ ]、多重比較の結果、問4は問5を除く全5問(問1, 2, 3, 6, 7)よりも有意に低く、問5も問1, 問2, 問7よりも有意に確率が低かった。非典型的誤答についても問題の効果が有意であり [ $F(6, 86) = 2.16, p < .05, \eta_p^2 = 0.03$ ]、多重比較の結果、問4は問1よりも有意に高いことが示された。

次に、通常教示条件における各参加者の確信度について平均と標準偏差を求め (Figure 2)、確信度に対して、問題(7)×回答の正否(2)の2要因参加者間計画の分散分析を行った結果、問題と正誤のいずれの主効果も有意であり [問題,  $F(6, 583) = 14.32, p < .0001, \eta_p^2 = 0.13$ ; 回答の正誤,  $F(1, 583) = 180.44, p < .0001, \eta_p^2 = 0.24$ ]、両者の交互作用も有意であった [ $F(6, 583) = 8.01, p < .0001, \eta_p^2 = 0.08$ ]。下位検定の結果、正答時の確信度は、問6は問1よりも低く、問7は問1, 問2, 問3よりも有意に低かった。一方、誤答時の確信度は、問1が問2, 問3, 問4, 問7よりも高く、問2が問3よりも高く、問5が問3よりも高く、問6が問3, 問4, 問7よりも高かった。また各問題における正誤間の効果量を比較すると、問2, 問3, 問4, 問5, 問7の $\eta_p^2$ が0.03から0.15までの間に分布する一方で、問1と問6については $\eta_p^2$ が0.009と0.007と正誤

Table 2  
各条件での設定とデモグラフィック変数の記述統計

教示	問題の呈示順序	男性	女性	合計	平均年齢	SD
通常	問1問2問3	9	6	15	19.13	0.99
通常	問1問3問2	7	7	14	19.29	2.13
通常	問2問1問3	11	4	15	19.33	1.88
通常	問2問3問1	9	5	14	18.86	1.10
通常	問3問2問1	10	5	15	18.73	0.70
通常	問3問1問2	8	5	14	19.08	0.95
時間	問1問2問3	12	8	20	18.85	0.88
正確	問1問2問3	13	7	21	18.95	0.83
全体		79	47	128	19.02	1.23

の効果が小さかった。この差は、問1と問6については、正答－誤答間での確信度の差が小さかったことを示している。

**問題の間の呈示順序の効果**

次に、通常教示条件における問1から問3の正答率を呈示順序ごとに求めた（Figure 3）。問題（3）×順序（3）の2要因参加者間分散分析を行ったところ、問題と順序の主効果、および両者の交互作用のいずれも有意でなかった〔問題,  $F(2, 252) = 0.47, p = .62, \eta_p^2 = 0.004$ ; 順序,  $F(2, 252) = 0.48, p = .62, \eta_p^2 = 0.004$ ; 交互作用,  $F(4, 252) = 1.44, p = .22, \eta_p^2 = 0.022$ ]。

次に、通常教示条件における参加者の確信度について、問題と呈示順序の効果を検討した（Figure 4）。問題（3）×順序（3）×正誤（2）の3要因参加者間計画の分散分析を行った結果、問題、呈示順序、正誤の主効果すべてが有意であった〔問題,  $F(2, 241) = 27.63, p < .0001, \eta_p^2 = 0.19$ ; 呈示順序,  $F(2, 241) = 3.62, p < .05, \eta_p^2 = 0.03$ ; 回答の正誤,  $F(1, 241) = 125.13, p < .0001, \eta_p^2 = 0.34$ ]。呈示順序1番の確信度は、呈示順序3番よりも有意に高かつ

た。交互作用については、問題と正誤のみが有意であり〔 $F(2, 241) = 21.71, p < .0001, \eta_p^2 = 0.15$ 〕、呈示順序との交互作用はいずれも有意ではなかった。下位検定の結果、問題の間における確信度の差は、正答時には有意でなく〔 $F(2, 241) = 2.17, p = .12, \eta_p^2 = 0.018$ 〕、誤答時には有意であった〔 $F(2, 241) = 34.94, p < .0001, \eta_p^2 = 0.23$ 〕。誤答時の確信度は、問1、問2、問3の順に高かった。

**教示の効果**

時間圧条件教示、通常教示（オリジナルの問題順序のみ）、そして正確条件教示のデータを対象として、問1から3の正答率を求めた（Figure 5）。教示の効果を検証するため、正答率に対して問題（3）×教示条件（3）の2要因混合計画の分散分析を行った結果、いずれの主効果、および両者の交互作用は有意ではなかった〔問題,  $F(2, 106) = 1.04, p = .36, \eta_p^2 = 0.019$ ; 教示条件,  $F(2, 53) = 0.37, p = .70, \eta_p^2 = 0.014$ ; 交互作用,  $F(4, 106) = 0.26, p = .90, \eta_p^2 = 0.01$ ]。

次に、確信度について、3つの教示条件ごとに、問1から3の平均確信度を求めた（Figure 6）。確

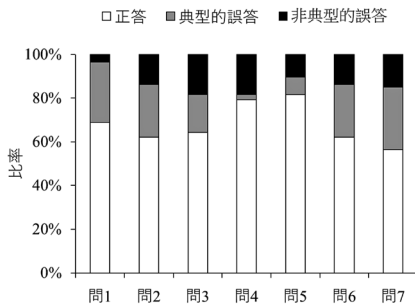


Figure 1. 通常教示条件の各問題における正答、典型的誤答、非典型的誤答の比率。

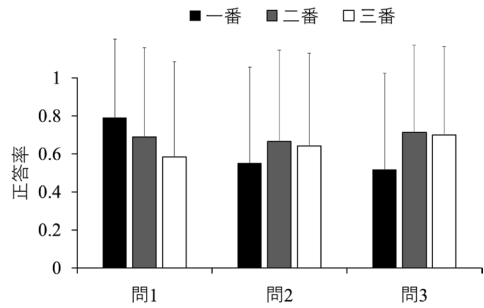


Figure 3. 通常教示条件における問1から3の呈示順序ごとの正答率。エラーバーは標準偏差を示す。

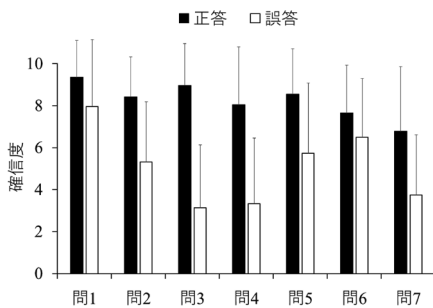


Figure 2. 通常教示条件の各問題における平均確信度。エラーバーは標準偏差を示す。

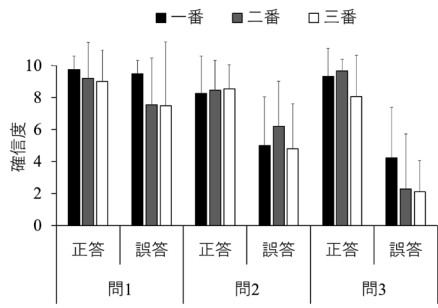


Figure 4. 通常教示条件における問1から問3の呈示順序ごとの平均確信度。エラーバーは標準偏差を示す。

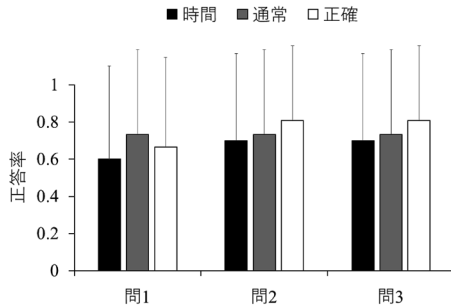


Figure 5. 時間教示条件, 通常教示条かつ通常順序, 正確教示条件における問1から3の正答率。エラーバーは標準偏差を示す。

信度に対して, 問題(3) × 教示条件(3) × 正誤(2)の3要因参加者間分散分析を行った結果, 問題と正誤の主効果は有意であった [ $F(2, 149) = 25.48, p < .0001, \eta_p^2 = 0.26$ ;  $F(1, 149) = 81.94, p < .0001, \eta_p^2 = 0.36$ ]が, 教示の主効果は有意でなかった [ $F(2, 149) = 0.13, p = .88, \eta_p^2 = 0.002$ ]。また, 交互作用についても, 問と正誤の間のみが有意であり [ $F(2, 149) = 10.89, p = .0001, \eta_p^2 = 0.13$ ], 教示条件との交互作用はいずれも有意ではなかった。下位検定の結果, 正誤間における確信度の差は, 問1は有意でなく [ $F(1, 149) = 2.67, p = .11, \eta_p^2 = 0.018$ ], 問2と問3は有意であった [ $F(1, 149) = 43.28, p < .0001, \eta_p^2 = 0.23$ ;  $F(1, 149) = 51.58, p < .0001, \eta_p^2 = 0.26$ ]。

## 考 察

本研究は, 認知的熟慮性をCRTによって測定し, 利用していくための基礎データとして, 回答種類とその確率, ならびに確信度を分析することで, (a) Frederick (2005) のオリジナル3項目と追加的に作成された4項目間の等価性の有無, (b) オリジナル3項目について問題呈示順序の効果, そして (c) 時間圧と正確さ追求を強調した教示の効果のそれぞれを検討した。

### オリジナル3問題と追加4問題の等価性の分析

問1から7までの正答, 典型的誤答, 非典型的誤答の発生確率を比較すると, 問4と問5は正答が多く, 典型的誤答が少なかった。これは, 本来CRTでの測定を目指していた認知的熟慮性と反する「システム1による直感的な判断」が生じにくい問題であることを示しており, 問4と5をCRTオリジナル項目の代替問題として用いることは妥当ではない

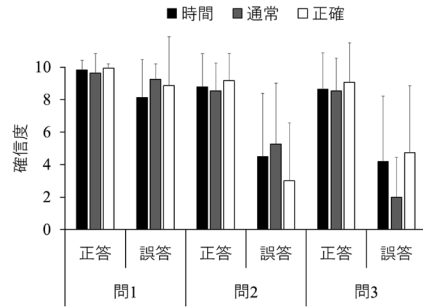


Figure 6. 時間教示条件, 通常教示条件, そして正確教示条件における各問題の平均確信度。エラーバーは標準偏差を示す。

と考えられる。

また, 正答時の確信度については, 問6, 問7でオリジナル項目よりも低く, とくに問7は, 正答を回答しているにもかかわらず, メタ認知では, その回答に対して自信を持っていないことから, 他の項目とは等価ではないと考えられる。その原因として2つの可能性が考えられる。1つは, 問7はオリジナル3問と比較して, 情報量が多いことである。特に不要な情報(例「2008年」, 「80万」, 「6か月後」)が多く, 正答するためには, これらの情報を「不要として思考の対象から外す」必要がある。その結果として, 正答に至るまでの認知的過程が複雑であり, 負荷が高くなっているために, 確信度を低下させていると考えられる。もう1点, 問7のみ, 回答を3肢強制選択から選択するようになっており, 必ずしも自分自身が主体的に答を産出せずとも答えられる, あるいは外的な手掛かりを基に反応することから, 他の問題のように, 回答者自身が直接的に計算・推定した場合の正解よりも, 回答に自信が持てない可能性が考えられる。

以上から, Toplak et al. (2014) で提案された4つの追加問題については, 少なくとも日本語使用者が日本文化の中で用いる場合には, 問4・5は回答比率の面から, また問7は正答時確信度の面から, オリジナル3問と等価な認知活動を測定できる問題とはいえない。追加項目の内, 問6については, 正答時の確信度がやや低いものの, 回答比率はオリジナル3問との間で差がなかったため, オリジナル項目と同様に認知的熟慮性を測定できる項目として利用可能性が高いと考えられる。

### オリジナル3問題の間の呈示順序の効果

回答比率の結果からは, 問題の間の呈示順序はオリジナル3項目の回答には影響を与えていないこと

が示された。その一方で、確信度は呈示順序1が最も高く、問題2、3と進むにつれて低くなることが示された。こうした結果の原因の1つは、CRT問題が一定の問題構造を持っており、回答者が複数の問題を解くプロセスにおいて、それに気づきやすくなることが考えられる。CRTが持つ問題構造の特徴は典型的誤答を引き出すように設計されていることであり、CRT開始時には、回答者はそのような構造の存在を知らないため、自分自身の回答に対し、高い確信度を報告するが、問題を解きすすめていくにつれて、特殊構造に対する気づきやその際のメタ認知の変化が発生し、自分の回答に対して、確信度を低く評価する傾向がある可能性が考えられる。

しかし加えて、本研究の確信度の結果から、呈示順序とは独立の、問題の間で相違があることが示唆された。特に誤答時の問1の確信度が高い点は、先行研究（Cueva et al., 2016; Pennycook et al., 2015）において、問1の正答率が他の2問よりも低かったこととの関係性が考えられ、興味深い。この結果はオリジナルCRTの中でも、問1は特異な問題である可能性を示唆している。この点を踏まえると、先行研究ではCRTの各問題に対する回答を1問1点として得点化することが多かったが、その数量化は必ずしも適切ではないと考えられる。一方で本研究では、正答率の差は問1から3の間で有意でなく、確信度の差のみが有意であったため、問1のどのような側面が問2・3と異なるのかについては明らかにされなかった。問1の特異性に関しては、今後さらに検討を続ける必要がある。

### 教示による時間圧／正確性強調の効果

回答比率ならびに確信度の両方において、時間圧操作ならびに正確性強調の教示による操作は有意な効果を示さなかった。このことは、CRTでの認知的熟慮性測度が外的な問題状況設定の変化によって変わりにくい可能性が示されたと考えられよう。

しかし、今回の研究では、紙と鉛筆による冊子を用いた集団実験であること、教示も「表紙に書かれた文」のみであるために、この一つの結果から、教示効果がないと結論付けることは危険であろう。実際、外的条件の効果を示した先行研究（Simonovic et al., 2017）では、スピーチ課題によるストレスを、ビデオカメラとマイクロフォンを向け、鏡とタイマーが見える環境下で「これからスピーチを行う」と通知された後にCRTに回答している。個人実験で、物理的にも外部環境を変更した状態で、さらに時間圧や正確さの追求が認知的熟慮性に及ぼす影響

について検討を行う必要があるといえよう。

### 結論

以上、オリジナルのCRTの代替問題としては、4問中1問のみが利用可能性があること、オリジナルの3問題の間にも特性の違いが考えられること、問題呈示順序には特に確信度において一定の効果があるが、回答比率には影響がないことが示された。また教示による影響、ならびに、時間圧-正確さ追求の変化が、認知的熟慮性判断に及ぼす効果については、さらに検討を行っていくべきであることが示された。

### 引用文献

- Campitelli, G., & Gerrans, P. (2014). Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach. *Memory & Cognition, 42*, 434-447.
- Cueva, C., Iturbe-Ormaetxe, I., Meta-Perez, E., Ponti, G., Sartarelli, M., Yu, H., & Zhukova, V. (2016). Cognitive (ir) reflection: New experimental evidence. *Journal of Behavioral and Experimental Economics, 64*, 81-93.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*, 25-42.
- Jacoby, L. L., Kelley, C. M., & McElree B. D. (1999). The role of cognitive control: Early selection versus late correction. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp.383-400). New York: Guilford Pub.
- Kahneman, D. (2003). A perspective on judgement and choice: Mapping bounded rationality. *American Psychologist, 58*, 697-720.
- Kahneman, D., & Tversky, A. (1982). On the study of statistical intuitions. *Cognition, 11*, 123-141.
- 松原和也・杉山 洋・村越琢磨・増田知尋・本田秀仁・和田有史 (2015). 高齢者の認知傾向とインターネットでの購買行動の関係 映像情報メディア学会誌, 69, 271-277.
- 岡本真彦 (1991). 発達要因としての知能及びメタ認知的知識が算数文章題の解決におよぼす影響 発達心理学研究, 2, 78-87.
- Pennycook, G., Cheyne, J. A., Koehler, D., & Fugelsang, J. A. (2015). Is the cognitive reflection test a measure of both reflection and intuition? *Behavior Research Methods, 48*, 341-



348.

佐藤安里沙・柴田みどり・伊藤友一・寺澤悠理・辻幸樹・田仲祐登・梅田 聡 (2017). 知覚課題と記憶課題における確信度 Type1・Type2信号検出理論を用いた生理心理学的検討 日本認知心理学会第15回大会発表論文集, 4-28.

Simonovic, B., Stuppel, E. J. N., Gale, M., & Sheffield, D. (2017). Stress and risky decision making: Cognitive reflection, emotional learning or both. *Journal of Behavioral Decision Making*, 30, 658-665.

篠塚寛美 (1993). 知的判断の確信度と正確度 心理学研究, 63, 396-403.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, 39, 1275-1289.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & Reasoning*, 20, 147-168.

Travers, E., Rolison, J. J., & Feeney, A. (2016). The time course of conflict on the cognitive reflection test. *Cognition*, 150, 109-118.

(受稿 4月27日 : 受理 5月29日)