

論文

批判的思考力テストの妥当性についての一考察： 久原他（1983）に焦点をあてて

Validity of the Critical Thinking Test Developed by Kuhara et. al. (1983)

峯島 道夫・今井 理恵
MINESHIMA Michio¹ and IMAI Rie²

This paper discusses the validity of the critical thinking skills test developed by Kuhara, Inoue, and Hatano (1983). Critical thinking is considered to be one of the 21st century competencies and recognized as of vital importance in general school education including English language teaching. Some academic studies, however, which use the critical thinking test developed by Kuhara et al. (1983) to validate their findings report statistically significant results without clarifying which construct of critical thinking they have focused on. This paper examines Kuhara's critical thinking test in detail and argues that since it is specifically designed to measure the inferential ability of critical thinking, it should be used for that purpose alone and not for any others.

キーワード： 批判的思考力、批判的思考力テスト、推論、状況モデル

Key words： critical thinking, critical thinking test, inference, situation model

1 はじめに

本論文は、久原・井上・波多野（1983）が作成した「批判的思考力テスト」を、英語教育における批判的思考（以下、CT）指導の有効性の検証のための使用にすることの妥当性について論じたものである。先行研究である峯島・今井（2021）では、同じ問題を端的に指摘することどめたが、今回は掘り下げて詳細に論じる。

CTは21世紀型コンピテンシーの一つとみなされ（P21, 2009）、学校教育においてもその重要性が一般に認知されるようになってきた（中央教育審議会, 1996；文科省, 2011, 2012a）。それ自体は好ましいことであるが、文部科学省（2012b）がCTを「多様な観点から考察する能力（クリティカル・シンキング）」のように単純にしか説明していないことも関係するのかもしれないが、「批判的思考概念が一枚岩であるという誤解」（道田, 2015）も広く認められるように思われる。しかし、CTは複数の構成要素から成る複合的な能力である。たとえばEnnis（1987）は、CTを態度（disposition）と能力（ability）の2つに大別し、態度に「他の可能性を探す（look for alternatives）」、「他者の感情、知識レベル、洗練さの度合いに敏感になる」など14項目を、能力に「演繹法」、「帰納法」、「価値判断をする」など12項目を含めている（pp. 12-13）。また、楠見（2011）は、CTに情報の明確化・情報の分析・推論・行動決定の4つのプロセスを認め、情報の明確化のプロセスには（a）問題、主題、仮説に焦点を当てて、それを明確化する、（b）情報の構造と内容を明確化する、（c）明確化のための問いを發する、（d）用語の定義や比喩や類推などの同定を行う、（e）隠れた前提を同定するといった構成要素があると述べている（pp. 8-9）。

このように、CTは複合的な能力であるにもかかわらず、CTを研究テーマとする実践・研究報告の中には、CTのどの構成要素を対象とし焦点化したのかが不明なまま、指導や活動によって学習者のCTが伸びたと主張するものがある（4節参照）。これはまるで、体育科教育において、ある指導を行ったら生徒の運動能力が伸びた、と主張するようなものである。伸ばそうとした運動能力が瞬発力なのか持久力なのか、筋力なのか敏捷性なのか、体の使い方なのか判断力なのかを明示しないのであれば、その効果測定に用いた方法が本当に適切であったかどうか、指導の効果があつたのかどうか議論できないはずである。同様に、英語教育の実践研究において、焦点化されたCTの構成要素が明示されないまま、久原ら（1983）のCTテストを効果検証に用いて効果があつたと主張することにも妥当性はないはずである。

本研究では、最初に久原ら（1983）のテストの分析を行い、同テストが推論に特化したテストであることを明らかにする。次に、英語教育において用いられる推論と同テストの推論の違いについて考察する。最後に、同テストを用いてCTの伸長に効果的であったとする先行研究の妥当性を批判的に検討する。本研究によって今後のCT研究の適正な発展に寄与したい。

2 久原ら（1983）の「批判的思考力テスト」

本節では、久原ら（1983）のCTテストの特徴を明らかにする。それによって、同テストが「推論」に特化したテストであること、さらに、3節で示す英語教育の読解において用いられる推論とは性質をかなり異にしていることを示す。

2-1 久原ら（1983）のCTの定義とワトソン・グレーザーの「推論」のテスト

久原ら（1983）は、批判的思考力を次のように定義している。

「文章を読んでその内容について解釈を構成する過程（すなわち読解過程）では、文章に明記されている内容と常識的な知識を使って、さまざまな推論がひき出される。推論によってひき出された情報は、その確かさの程度が評価され、それに応じて文章の解釈を構成する際に利用される、と考えた。そしてこの過程で、推論の確かさの微妙な区別をする能力を批判的思考と名づけた。」（p. 137）

すなわち、久原ら（1983）は、文章の読解過程において引き出されるさまざまな推論の確かさの微妙な区別をする能力をCTと定義したのである。この定義は、文部科学省（2012b）の「多様な観点から考察する能力」という、文章の読解には全く限定されない広義の定義に対して、CTの一構成要素に限定した狭義の定義である。

久原らのCTテストは、完全なオリジナルではなく、作成にあたって参考にした原本がある。それは、Watson-Glaser Critical Thinking Appraisal（1964）（以下、WGテスト）である。久原らは、CTと定義した推論の能力の測定に適したテストを、当時のアメリカで使用されていたいくつかのCTテストから探していたが、このWGテストの下位テストのひとつが適当であると判断したと述べている（p. 133）。WGテストは次の5つの下位能力から成っている。それは、①inference（推論）、②recognition of assumptions（前提の同定）、③deduction（演繹的推論）、④interpretation（解釈）、⑤evaluation of argument（議論の評価）であり、久原らが選んだのは①推論であった。WGテストにおけるこの①推論を、Norris & Ennis（1989）は次のように説明している。

The Inference subtest is designed to test whether the examinee has the ability for judging the likelihood, given certain evidence, of a conclusion being true or false.

(下位テストの一つである「推論」は、受験者が、ある証拠が与えられたとき、[推論によってもたらされた]結論がどの程度真または偽でありそうかを判断する能力を持っているかどうかをテストするように作られている。 (*訳は著者による。以下すべて同じ。)

ここで述べられている推論のテストが意図している能力と、久原ら (1983) の定義にある「…[文章の読解]過程で、推論の確かさの微妙な区別をする能力」は一致していることが分かる。後の議論を少し先取りすれば、久原らのCTテストとその原本のWGテストの推論は、英語教育において用いられる推論とは性質を異にしている。英語教育における推論は、テキストのより深い理解 (状況モデルの構築) に必要とされるから行われる (逆に言えば、必要のない推論は行われない) のに対して、久原らの推論は、読む目的とは無関係に行われる。次に、久原ら (1983) のCTテストを詳細に分析し、推論をどのようにとらえていたかを検証する。

2-2 推論の4タイプと5段階の判定基準

(1) 推論の4タイプ

久原らのCTテストがWGテストを原本としていることは先に述べたが、作成にあたって久原らはWGテストを批判的に考察し、WGテストでは推論の種類が明示されていないこと、中には推論と言えない単なる言い換えに過ぎないものが含まれていることを考慮し、推論のタイプを4種類に設定するという改定を加えたとのことである (p. 133)。次はその4種類の推論の名称と簡単な説明である。

- (1) タイプ1 : クラスの形式的推論 (formal class reasoning) — 「個々の要素について真なら、全体についての全称否定は偽であり、全体について真なら、その要素についても真であるという推論」
- (2) タイプ2 : 帰納を含むクラスファジー推論 (fuzzy class reasoning) — 「いくつもの要素について真であれば、全体についてもたぶん真であり、全体について、はっきりした傾向があれば、要素についても、その傾向はたぶんあるという推論」
- (3) タイプ3 : 省略のあるファジー条件文推論 (abbreviated fuzzy conditional reasoning) — 「大前提を常識的知識で補うと、条件文推理の形で導き出される推論」
- (4) タイプ4 : 前提の認定 (recognition of presupposition) — 「与えられた情報の前提となっていることからの認定の推論」

また、同テストには、SM (2) とTM (2) という、形式と推論のタイプは同じであるが、内容が異なる2つの平行テストがある。(同テストをCTの指導の効果測定に用いる研究は、2つのテストをプリテストとポストテストとして使っている。) どちらのテストも4つの大問 (含まれる小問は計20問) から成り、1つの大問は、約200字から450字の長さの、真であるとされる1つの問題文と確かさの程度の判定が求められる4文から7文の推論の文で構成されている。

(2) 5段階の判定基準

久原らのCTテストの受験生は、出題された推論の文の確かさを、次の5つの基準により判定す

る。

- ①真：「推論が全く正しいと思われる場合」
- ②たぶん真：「5割以上の確かさで正しいと思われるが、絶対に真とは断定できない場合」
- ③材料不足：「判断の材料が不十分だと思われる場合」
- ④たぶん偽：「5割以上の確かさで誤っていると思われるが、絶対に偽とは断言できない場合」
- ⑤偽：「推論が全く誤りだと思われる場合」

この5段階の判定が「推論の確かさの微妙な区別をする能力」の具体的な中身である。同テストが実際にどのようなものなのか、さらに、同テストの推論が後に論じる英語教育における推論発問の推論とどのように違うのかを明らかにするために、同テストの問題例の1つを次に示す。

<ベン・カーターの新聞>

ベン・カーターはA州最初の新聞の編集者である。彼は父の経営する印刷所を手伝いながら、独学した。彼が編集した最初の新聞は、1710年9月25日A州の首府で発行されたが、その日のうちに州知事によって発行を禁止された。ベン・カーターはこれに屈せず、なおも自分の小さな新聞を発行しつづけ、書きたい記事を書くために長いたたかいを続けた。これは今日も続いている出版の自由をまもるための努力のなかで、ひとつの重要なエピソードになっている。

推論1：A州知事は、ベン・カーターの新聞の発行禁止処置は知事の権限内にあると考えた。

推論2：ベン・カーターは、自分の興味や目的にとりついたら離れない頑固な人間だった。

推論3：1710年ころのA州では、言論の自由が全面的には認められていなかった。

推論4：この最初の新聞には、A州知事を攻撃する記事が書かれていた。

推論5：ベン・カーターは、彼が編集した新聞のために、その後も幾度も罰金を払わねばならなかった。

この問題文は4つの大問の中で最も短いものである。推論1はNuttall（1996, pp. 188-189）の発問の分類に従えば、Type 2の Questions involving reorganization or reinterpretation（情報の再編成または再解釈）に属するだろうが、久原ら（1983）の解説によれば「前提の認定」であり、「州知事によって発行を禁止された」ということは、それが知事の権限内にあったからだと推論できれば良い。推論2は「クラスのファジー推論」の例で、「おそらく真」が正解とのことである。推論3は「クラスの形式的推論」の例であり、「真」。推論4と5はテキストに書かれていないので「材料不足」である。

久原ら（1983）は、これらの推論の確かさの微妙な区別をする能力を読解における批判的思考力と定義した。しかし、著者の個人的な感覚では、たとえば推論2の「ベン・カーターは、自分の興味や目的にとりついたら離れない頑固な人間だった。」という命題の確かさが「真」ではなく「おそらく真」だと判別できる能力が、テキストの読解において重要であるとも、批判的思考力であるとも思えない。実際このテキストを読んだ時に、彼はきっと信念のある人だったのだろうとは考えたが、推論2のようなことは全く考えなかった。問題として出題されたからその確かさを事後的に熟考したに過ぎない。同テストが「批判的思考力」と定義した、推論の確かさの微妙な区別をする能力の重要性については疑問が残る。

3 英語教育の推論と久原ら (1983) の推論の違い

本節では、久原ら (1983) のCTテストの推論と、日本の英語教育の読解において扱われる推論の違いについて考察する。これは、後出の、指導によってCT力が向上したと主張する実践研究の妥当性の検証にも役立つはずである。

3-1 英語教育の読解における「推論」

英語の読解に関して、卯城 (2009) は「内容理解を伴った読解では、適切な状況モデルを構築することこそが最終目標であり、また、その状況モデルを構築するにあたって読み手は、自分の背景知識を活性化させ、そして推論を働かせなくてはなりません。」(p. 85) と述べている。状況モデルとは、読解によって生成される心的表象 (mental representation) の3つのレベルである、表層的記憶 (surface memory)、命題的テキストベース (propositional textbase)、状況モデル (situation model) の最後のレベルの心的表象で、「文章からの情報と読み手がそれ以前に持っていた背景知識が統合された、テキストによって記述される状況全体の理解のこと」(p. 84) を指す。ここで重要なことは、英語教育における読解指導の最終目標が、意味相互に矛盾がなく、首尾一貫した状況モデルを構築することであり、そのために明示されたテキスト情報と自らの背景知識を活用し推論を働かせることが必要になる、ということである。逆に言えば、英語教育における推論は、テキストの状況モデルの構築という最終目標に貢献する限り、という条件の下に行われ、状況モデルの構築に不必要な推論は行われまいということである。

卯城 (2009) と同様の主張は、Grabe (2009) にも見出せる。Grabe (2009) は命題的テキストベースではなくテキストモデル (text model) という用語を用いているが、2つは同義であり、レベル2と3の心的表象について次のように説明している。

The requirements for generating an appropriate situation model are quite distinct from the requirements for a text model. The text model calls for understanding what the text itself is trying to signal. The situation model calls for the reader to combine background knowledge with text information and assists in an effective interpretation of the text in line with the reader's goal. (p. 46)

(適切な状況モデルの創造に必要な要件は、テキストモデルに必要な要件とははっきり異なる。テキストモデルは、テキストそれ自体が伝えようとしていることの実理解が必要とされる。状況モデルは、読み手に背景知識をテキストの情報と組み合わせることを求め、読み手の目的に沿ったテキストの効果的な解釈を助ける。)

卯城 (2009) は、読解の最終目標は状況モデルの構築であると述べているが、上記の説明でもあるように、状況モデルは、新たに加わるテキストの情報の解釈をより妥当性の高い方向へと導くと同時に、それ自体も新たに加わるテキストの情報によって、必要が生じれば修正や補強を受けつつ構築されるものである。さらに、読み手は、ただ無目的に漫然とテキストを読むのではなく、自らの読む目的に沿って読み進め、背景知識とテキスト情報を活用しつつ推論を行っているのである。

先に示した久原らのCTテストの〈ベン・カーターの新聞〉の問題は、この状況モデルの構築という観点からすると、推論3 (「1710年ころのA州では、言論の自由が全面的には認められていなかった。」) は状況モデルの構築に貢献すると考えられるが、それ以外の推論はおそらく必要ない。

田中 (2011) は、リーディングの指導で教師が学習者に投げかける発問には、①事実発問 (fact-

finding questions)、②推論発問 (inferential questions)、③評価発問 (evaluative questions) の3つがあり、②の推論発問を「テキスト上の情報をもとに、テキスト上には直接示されていない内容を推測させる」発問のことであると定義している。推論発問の具体例として次の例を挙げ、解説を続けている。

例 文：Mary flies to the Occident twice a year to buy fashionable clothes.

推論発問：Does Mary have a lot of money?

解 説：Maryが裕福かどうかはテキスト上に直接書かれていません。しかし、“Mary flies to…”という部分や、“to buy fashionable clothes”という情報を読み、「Maryは飛行機で年に2回、しかも、流行の服を買うためにどこかに行く」という情報から、Maryは恐らく裕福であることが推測できます。」(p. 13)

この例の推論発問の目的は、学習者がMaryについての命題的テキストベースの理解を超えて、より充実した状況モデルを構築する手助けをすることであろう。このような推論発問は、英語教育の内容理解の活動において広く一般に用いられる発問である。

Graesser, Singer, and Trabasso (1994) は、物語文における知識基盤の推論 (knowledge-based inference) には、次の13種類 (class) があるとする。すなわち、①Referential (照応関係)、②Case structure role assignment (格意味役割付与)、③Causal antecedent (因果的先行詞)、④Superordinate goal (上位のゴール)、⑤Thematic (主題)、⑥Character emotional reaction (登場人物の感情的反応)、⑦Causal consequence (因果的結末)、⑧Instantiation of noun category (名詞カテゴリーの例示)、⑨Instrument (道具)、⑩Subordinate goal-action (下位のゴール/行為)、⑪State (状態)、⑫Emotion of reader (読み手の感情)、⑬Author's intent (著者の意図) である。これらの推論のタイプが具体的にどのようなものなのかを示すために、同論文で取り上げられた元になるテキストと推論の例を示す。

“How Leisure Came” by Ambrose Bierce

A Man to Whom Time was Money, and who was bolting his breakfast in order to catch a train, had leaned his newspaper against the sugar bowl and was reading as he ate. In his haste and abstraction he stuck a pickle-fork into his right eye, and on removing the fork the eye came with it. In buying spectacles the needless outlay for the right lens soon reduced him to poverty, and the Man to Whom Time was Money had to sustain life by fishing from the end of the wharf. (p. 375)

同論文は、推論の13タイプと、各タイプの推論をこのテキストに当てはめた場合に導かれる推論の結果を示しているが、下の表1は、それらに加えて、可能な推論発問(表中央の列)を提示したものである。

表1. 推論のタイプと推論発問

Type of inference	Possible inferential question	Inferences
1. Referential	What does "it" in "...on removing the fork the eye came with it" refer to?	Fork is the referent for it.
2. Case structure role assignment	Where did he put his newspaper?	Against the sugar bowl is assigned to a location role.
3. Causal antecedent	What caused him to stick a fork into his eye?	The man was careless and mis-aimed his fork.
4. Superordinate goal	Why was he in such a hurry to catch a train?	The man wanted to get to work and earn money.
5. Thematic	What is the main point or moral of the text?	Haste makes waste.
6. Character emotional reaction	How did he feel after the accident?	The man became sad.
7. Causal consequence	After his right eye came with the fork, what happened to him physically?	The man became blind in his right eye.
8. Instantiation of noun category	What do you think he had for breakfast?	Bacon and egg.
9. Instrument	What do you think he used when he fished?	The man used a rod and reel (to fish).
10. Subordinate goal-action	How did he eat his breakfast?	The man grasped his fork and moved it toward his mouth.
11. State	What did the city he lived in have?	The city has a wharf.
12. Emotion of reader	How do you think the reader feels when he/she reads "on removing the fork the eye came with it"?	The reader is disgusted.
13. Author's intent	What do you think motives the author to write this piece?	Bierce wants to lambaste workaholics.

(Graesser, et.al. 1994, p. 375 を参考に作成)

このうち、①の代名詞とそれが指し示す名詞の照応関係であるReferential（照応）と、②の主格、与格、目的格、前置詞等によって表されるCase structure role assignment（格構造役割割当）は、文法上の知識があれば自明のことである。（①はいわゆる橋渡し推論である。）また、⑫の読み手の感情は、テキストの読解とは別の推論となる。（むしろ、著者が読者の感情的な反応をどのように予想して書いたのかを、読者自身が作品の評価のために推論することならばあり得るので、⑬に含めた方が良くとも考えられる。）したがって、これらを推論と呼ぶべきかどうか疑問の余地が残るが、それ以外の推論は、日本での英語の授業においてもテキストをより深く読むために使われる、普通の推論発問だと考えられる。これらの推論発問に答えられるようになることが、先にみた久原らのCTテストの推論—「A州知事は、ベン・カーターの新聞の発行禁止処置は知事の権限内にあると考えた。」や「この最初の新聞には、A州知事を攻撃する記事が書かれていた。」—の確かさの程度を正しく査定する能力の伸長に繋がるかは疑問である。

3-2 久原ら（1983）のCTテストの「推論」の問題点

繰り返しになるが、久原ら（1983）は、読解過程において「推論によってひき出された情報は、その確かさの程度が評価され、それに応じて文章の解釈を構成する際に利用される」と考えた。換言すれば、読者は、さまざまな推論をひき出し、それらの推論の確かさの程度に応じて、文章解釈に利用する、との主張である。しかしこの考え方は妥当だろうか。著者自身の個人的な読みの実感に照らすと、読解において何らかの推論をしている時は、いくつかの可能な推論を考え、次にそれら複数の推論について、その確かさの程度を真、たぶん真、たぶん偽、偽、判断材料不足、などとレベル分けなど全くしていない。そのような読み方は非効率的で意味がないだろう。むしろ、推論するときは、それまでに構築した状況モデルと最も矛盾なく一貫性を保つような「おそらく真」の推論のみを考えるはずである。おそらく偽や偽の推論をする必要など全くない。本稿 2-2 で久原らのCTテストの一例をみたが、推論 4 や 5 を読み手自らが考えることはない。久原（1983）は、CTを文章の読解過程において引き出されるさまざまな推論の確かさの微妙な区別をする能力と定義づけたが、英語教育においては、確かさの異なるいくつかの推論の微妙な区別など必要ではなく、必要とされるのは、最も真だと思われる推論だけである。

自然な読みにおいては、読み手はそれまでの状況モデルと最も整合性の高い推論を無意識のうちに生成するだろう。またその推論も、現時点で最も真である可能性が高いが、新しい情報と矛盾した場合には修正すればよい、融通の利く推論である。たとえば、Sanfort & Garrod（1981）の次の例におけるJohnは何者だろうか。

John was on his way to school. ①He was terribly worried about the mathematics lesson. ②He thought he might not be able to control the class again today. ③He thought it was unfair of the instructor to make him supervise the class for a second time. ④After all, it was not a normal part of a janitor's duties.

(p. 132；文中の○番号は著者による)

①までの文を読んで、学校へ行くのは普通生徒であるし、数学の授業を心配している点から考えても、おそらくJohn = studentなのだろうと推論するのが自然だろう。ここで他の可能性、たとえば、Johnはある生徒の父親で、息子が前から先生が嫌いだと言っていた数学の授業で何か問題を起こしたので、呼び出しを受けて学校へ向かっているのかもしれない、などという推論の可能性については考える必要はない。あるいは、教育実習生であったり、給食の配達員であったりする可能性についても考える必要はない。とりあえず、John = studentという推論で十分で、現時点で最もありそうな想定をしていれば読解に支障はないからである。しかし、②の文の授業をコントロールする、という新たな情報を受けて、John = studentの推論は否定され、John = teacherという新たな推論が打ち立てられる。しかし、③の文によって、その推論も否定され、④でJohn = janitor が判明する。

この例から明らかなように、読み手は、久原ら（1983）が示唆するように、確かさの異なる複数の推論の候補を同時に生成し、それらの確かさの差異を区別などしていない。読み手は、推論を構成する段階で、おそらく真であると思われるものを常に選択しているからである。推論に複数の可能性（alternative）が生じる場合もちろんあるが、それでも、それらはすべて真である可能性を持っているからである。

次に、久原らのCTテストの問題の一つを取り上げて分析し、英語教育における読みの最終目標である状況モデルの構築という観点から考察を加える。

＜A国軍極寒訓練＞

1946年A国軍はスノードロップ作戦とよばれる大規模な実験を行った。この実験の目的は、北極の厳しい気象条件のもとでどのような軍人が一番良く任務を果たしうるかを調べることである。全国の10の基地からこの実験への参加者を集めた。このなかには北欧系民族の血をひく者も、ラテン系民族の血をひく者もいた。頑丈な者もきゃしゃな者もいた。命令によって参加した者もいれば、志願した者もいた。血圧は、正常な者も、少し高めのも、低めのもいた。この作戦参加者は、全員、極寒の中で生存し、任務をはたすための訓練を、前もって受けた。この実験によって得られた資料は、コンピュータで数量的に処理された。その結果、北極における作戦において、効果的な行為をする人々と、効果的でない行為をする人々のあいだに差があるとされた要因は、調査されたさまざまな要因のうちのふたつだけであった。その(1)は、自らすすんで参加したかどうか(志願した者対命令によって参加した者)であり、その(2)は、北極という条件のもとで自分の身を守るための、知識と技術の程度である。

(1) 推論1：きゃしゃな者だけについてみると、自らすすんで参加した者のほうが、そうでないものよりも任務を遂行したとはいえない。

この推論の判定に関係する文は、「頑丈な者もきゃしゃな者もいた。」の文と、優劣を決める要因の一つが「自らすすんで参加したかどうか(志願した者対命令によって参加した者)」の2文である。推論の前半に書かれている「きゃしゃな者だけについてみると」という限定条件をどう理解すればいいのだろうか。テキストには参加者の特徴別の集団の調査結果など記述されていないので、「材料不足」と考えるのが妥当なのだろうか。久原ら(1983)によれば、この推論は「おそらく偽」で、推論のタイプは2の「クラスのファジー推論」とのことである。しかし、この推論の確かさを正しく判定できるかどうかは、「クラスのファジー推論」という集合論の概念を知っているか否かにかかっている。一度この概念を理解すれば、同様の問題にはすべて正解できるようになるからである。

この推論1は、テキストの状況モデルの構築には明らかに不要であると思われる。なぜなら、任務の遂行の適否を決める要因は、志願の形態とサバイバル知識と技術の習得の2つだけであり、体格的な特徴に関する情報はすべて周辺的であり忘れても構わないからである。この推論がおそらく偽だと判別できたとしても、このテキストを深く理解することにはならないだろう。

(2) 推論2：北極のきびしい気象条件のもとでは、もっとも効率よく任務を果たせるのは、北欧系民族だったとはいえない。

任務の遂行の適否を決める要因は、志願の形態とサバイバル知識と技術の習得の2つだけなので、この推論3は「偽」となり、それを否定しているので、推論3全体は「真」となるはずである。同論文によれば、この推論は、タイプ1の「クラスの形式的推論」である。Nuttall(1996, pp. 188-189)の発問の分類に従えば、おそらくこれはType 2の Questions involving reorganization or reinterpretation(情報の再編成または再解釈)に属し、推論には属さないだろう。この推論も、先の推論と同様に、肝心の実験結果とは無関係の情報についての内容であり、テキストの状況モデルの構築には寄与しないだろう。

(3) 推論3：A国軍では、当時、北極において軍事作戦を行うことがあり得ると想定していなかつ

た。

同論文によれば、これもタイプ3の「省略のあるファジー条件推論」である。この場合、たとえば、「もし北極において軍事作戦を行うのであれば、軍人は厳しい環境のもとでも効果的な行為を行う必要があり、そのためには軍は事前にどのような要因がそれを可能にするのかを実験によって知ろうとするだろう。」という条件文が考えられる。したがって、後件にある「実験を行った」という結果があるのであれば、それは「北極において軍事作戦を行う」ことが前件として想定されている（大前提とされている）と考えられる。

この推論も英語教育における読みであれば推論の対象にはなっていないだろう。なぜなら、もしも北極における軍事作戦が想定されていなかったら、そもそもこのような実験は行われるはずがなく、この当然のことはテキストを深く読むことに何の貢献ももたらさないからである。

(4) 推論4：南極での同様な作戦について調べたら、自らすすんで参加する者とそうでない者とは、任務に効果的な行為をするかどうかに関して異なるという結果は得られないだろう。

回りくどい表現の分かりにくさは別として、この命題は端的に言うと、状況が北極から南極に変わった場合、結果はどうなるかを尋ねている。同論文によれば、これは推論のタイプ2の「クラスファジー推論」の例ということである。先の推論2では、きゃしゃな軍人は全軍人を構成する一要素（部分）であると了解できるが、舞台が北極から同じ極寒の地の南極に変わった場合の推論も、これと同種の全体と部分の関係性に関わる推論だと考えるのは難しいかもしれない。

この推論は英語教育における読みでは出てこないだろう。なぜならテキストを理解するために、場面を北極から南極に変える必要など全くないからである。先に見たGraesser, et.al. (1994)の英語教育において用いられる推論発問にいくら答えられるようになっても、この推論4のような推論に正しく答えられる保証はない。適切な考え方を知っているかどうかだけが問題だからである。

(5) 推論5：A国軍では、この実験の参加者の各々が、極寒の中で自分の身を守るための知識と技術をどの程度もっているかを示す資料をもっていなかった。

本文には、効果的な行為をする人々とそうでない人々を差別化する要因の1つとして、「北極という条件のもとで自分の身を守るための、知識と技術の程度」が挙げられていた。すなわち、その知識と技術がどの程度だったのかという資料がなければ、問題の推論は下せない。したがって「偽」となる。

この推論も英語教育における読みでは出てこないだろう。なぜならこの推論の内容は、テキストに明示されている内容と明らかに矛盾しており、推論の必要性もないし、状況モデルの構築にも寄与しないからである。

3-3 まとめ

以上、みてきたように、久原ら（1983）のCTテストにおいて出題されている推論は、英語教育で用いられる推論とは性質を異にするものが多くある。英語教育において使われる推論は、テキストの内容をより深く読み、状況モデルを構築するために問われる必要性和必然性がある。無駄な推論は問われない。しかも、状況モデル自体が解答のヒントとなる。一方、久原らのCTテストの推論の中には、状況モデルの構築には貢献せず、その推論が問われる必然性が必ずしも保証されていないものが含まれている。さらに、先の推論1や2の問題のように、「クラスの形式的推論」や「ク

ラスのファジー推論」といった概念自体を知らないと正解が難しいものもある。また、推論6の「…異なるという結果は得られないだろう。」のように、回りくどい表現（「…同じ結果が得られるだろう。」となぜ表現しないのか）の使用もある。久原らは、推論の確かさの微妙な区別をする能力をCT力と定義したが、英語教育の読解においては、久原らのCTテストが求めるような種類の推論の識別力は重要だとは思えない。

4 CT研究と測定方法

本節では、CTを対象とした実践研究のうち、久原ら（1983）のCTテストをその指導の効果測定に用いたいくつかの事例を取り上げ、その妥当性について検証する。

4-1 研究目的と測定方法の整合性

平山（2004）は、CTが多く構成要素から成る複合的な能力であることから生じている、定義や研究成果における研究間の混乱について、次のように述べている。

批判的思考は、仮説、視点、疑問、解決、プランなどの生成を含んでおり、規準（criteria）に従う目標志向的な思考とも言われ（楠見, 1996）、多様な側面を持つため、定義も研究の目的により変化し、一致していないものと考えられる。また、中には定義を明示せずに行われている研究もある。多様な側面を持つため、一つの研究で網羅することはできないとすれば、その研究の中で検討されている側面は一体どのような側面であるのかを明示しなければ、研究の方法、結果、得られた知見等において、研究間で混乱を招くことになると考えられる。（p. 291）

すなわち、CTの実践研究においては、CTのどの側面、どの構成要素に焦点化したのかを明確にしないと、結果の妥当性が議論できないということである。久原らのCTの原本であるワトソン・グレーザーの批判的思考力テストが推論、前提の同定、演繹的推論、解釈、議論の評価の5セクションに分かれているのも、複合的な能力であるCTをその構成要素別に測る必要性があると考えているからである。平山（2004）は、さらにCT研究の目的と課題に関連して次のようにも述べている。

用いられる課題によって、測定される側面は大きく異なってくる。したがって、研究を行う際には、その目的に合った課題を注意して選ぶ必要がある。また、その目的と課題との一貫性を、その研究中に明示する必要がある。（p. 295）

この指摘も当然のことであるが、複合的な能力であるCTの力を評価しようとするならば、＜研究目的＞＜研究対象とするCTの構成要素＞＜その要素を測れる課題＞の3つが整合的に対応していることが必要である。先に、CTの定義を明示しない研究への言及があったが、CTを読みに適したクリティカル・リーディング（CR）の研究で、その定義を「読解後に英文の内容について意見を記述すること」としたものがある（伊佐地, 2014, p. 288）。しかし、読んだ英文について意見を記述するだけで、それが批判的思考やクリティカル・リーディングになる保証は全くない。その意見に含まれる主張、データ、論拠などを検証して初めて判断できるものである。主張の妥当性を吟味することこそCRであろう。

この平山（2004）の注意点を念頭におき、次に、久原ら（1983）のCTテストを、指導効果の検

証に用いている実践研究の例を取り上げ、そこで＜実践の研究目的・目標＞＜CTの定義と構成要素の特定＞＜課題（＝久原ら（1983）のCT（推論）テスト）＞の3つが整合的に対応しているのかを検証する。

4-2 事例1（伊佐地, 2016）

（1）実践の研究目的・目標

伊佐地（2016）は、自身が行った実践研究の「授業の目標」を次のように掲げている。「読んだ英文についてその主張や根拠などを検討し、自分の意見を論理的に構築すること、すなわちクリティカル・リーディングを実践し、英語力の向上とともにクリティカル・シンキングを学ぶことを目標とする。」（p. 36）すなわち、この実践の目標は、テキストの主張や根拠の分析と、それについて自分の意見の論理的な構築の2つをとおして、英語力を向上させCTを学ぶことである。

（2）実践の概要および推論の指導

本実践は、大学3年生9人を対象に、4月から7月にかけての15回の授業で行われたとのことである。久原らのCTテストを、指導効果の検証のために第1回の授業時に事前テストとして、第15回の最終授業時に事後テストとして行っている。実践の内容は次のようにまとめられている。「まず事実発問と推論発問によりテキストを深く理解させたのち、テキストの記述に対して疑問を投げかけたり、英文とは別の観点から考えさせたりするための発問（critical reading questions）に解答させたり、学生自身にこのタイプの発問を作らせたりすることで、英文の内容について批判的に評価させた。その結果に基づいて、テキストの内容について自分自身の意見を構築し意見文（評価発問）にまとめさせた。」（p. 55）

推論の指導に関して、While-readingの段階において「2巡目の読解（second reading）では、「推論発問」（例：Is the author for or against beauty contest?）を中心に、テキストの記述に基づいて推論し理解を深めるための発問を与えた。」（p. 36）との記述がある。これは本稿3-1で論じた、英語教育の読解において普通に使われる推論発問であり、Graesser et al.（1994）の推論の種類では「著者の意図」となるだろう。また、この他の推論発問の例として、“Does Ms. A completely hate his father?” “What is the hierarchical order of things?” “According to Mr. Fujiwara, why should Ms. A obey her father?” があげられているが、これらの推論発問は、いずれも内容理解のための、すなわち状況モデルの構築のための標準的な推論発問であり、久原らのCTテストで要求されるクラスの形式的推論やクラスのパジャー推論ではない。

（3）CTの定義と構成要素の特定

同実践に、道田・宮元（1999）の「何事も無批判に信じ込んでしまうのではなく、問題点を探し出して批判し、判断しようとする」とのCTの定義の引用はあるが、具体的にCTのどの側面に焦点を当てているのかは明示されていない。推論であるとも特定されていない。

（4）効果についての結論

同実践によれば、久原ら（1983）のCTテストを事前・事後に用いて平均値の差をウィルコクソンの符号付順位和検定による検討した結果、平均値の差は有意傾向を示し、大きな効果量も認められたため、「本授業実践は学生の批判的思考力の向上に一定の効果があったと言える」との結論を下している。

（5）まとめ

同実践は、読んだ英文についてその主張や根拠などを検討し、自分の意見を論理的に構築することを目標としていた。指導でCTの構成要素のどれに焦点化したのかは明示されていなかったが、

授業の目標に照らせば、WGテストの下位テストのひとつである evaluation of arguments (議論の評価) と、value judgment (価値判断) が適切だったであろうと思われる。平山 (2004) が指摘した、〈研究目的〉〈研究対象とするCTの構成要素〉〈その要素を測れる課題 (= 久原らのCTテスト) 〉の3つが整合的に対応しているとは言えない。

4-3 事例2 (有嶋ら, 2010)

(1) 実践の研究目的・目標

有嶋ら (2010) によれば、「本研究は、授業に英語ディベートを取り入れ、生徒の英語スピーキング力および批判的思考力の育成をめざした英語の授業を実践し、その効果を立証することを目標とした。」とのことである。すなわち、ディベート活動によってCTを伸ばすことが本実践の目標であった。

(2) 実践の概要および推論の指導

本実践では、普通高校の2年生1クラス (39人) を対象に、計12時間をかけて、ディベートの準備 (9時間) と試合 (3時間) を行ったとのことである。実践において推論をどのように指導したかについての言及はない。

(3) CTの定義と構成要素の特定

CTについては、「何を信じたり行ったりするかを決めるための、合理的で省察的な思考」(Ennis, 1987)、「日常生活において私たちが行う意思決定・判断・問題解決のために必要な認知能力と態度」(Facione, 1990) を引用し、さらに、「批判的思考は能力 (認知的側面) と態度 (情緒的側面) から構成される。具体的には、「能力」として、問題などの明確化、根拠に基づく判断、推論などが挙げられ… (以下省略)」との説明がある。しかし、同実践がCTのどの側面を焦点化したのかは明記されていない。推論の力とも特定されていない。

(4) 効果についての結論

久原ら (1983) のCTテストを事前・事後に用いているのは先の伊佐地 (2016) と同様であるが、違うのは、本実践では、「ディベートやスピーキングの特色を考慮して、問題文を読み返すことができない聴覚条件で実施した。問題文は1回、推論の部分は2回口頭で提示し、メモは取ってもよいこととした」点である。久原のCTテストは、本稿3-2で問題例を取り上げ、分析を加えたが、問題文のテキストは情報量が多く、メモを取ったとしても一度聞いただけではその内容を十分に把握するのは困難だろう。検証結果を分析するにしても、低い点数の原因が推論の力の欠如なのか、記憶の保持に困難があったからなのか特定できない。また、同実践では、久原らのCTテストをディベート活動 (2月) の前に2回 (9月と12月)、後に2回 (2月と3月)、計4回行っている。ディベート活動に最も近い前後の時期 (12月と2月) に行われた久原らのCTテストの全体の平均値は20点満点でそれぞれ11.4点と11.9点であった。LSD法による多重比較を行ったところ、プリテスト①とポストテスト②、プリテスト②とポストテスト②の差が有意であったとのことである。しかし、本来最も大きな差が表れるはずの、ディベート活動直前のプリテスト②と直後のポストテスト①の間には、有意な差は認められなかったことになる。ただし、全体を上位群と下位群に分けた結果では、下位群において2つの時期の差が統計的に有意であったと伝えている。

(5) まとめ

同実践では、実践の中核であるディベート活動によって、CTのどの構成要素を伸ばそうとしたのかが不明であった。また、ディベート活動の準備段階においても、意見整理、肯定・否定側の立論作成、立論の校正、質疑応答準備などは行われたが、文章の読解において推論を行うような活動

はなかった。したがって、＜研究目的＞＜研究対象とするCTの構成要素＞＜その要素を測れる課題（＝久原らのCTテスト）＞の3つが整合的に対応していたとは言えない。本実践の活動の効果検証には、文章読解における推論のスキルに特化した久原らのCTテストではなく、上記の事例1と同様に、evaluation of arguments（議論の評価）とvalue judgment（価値判断）を想定したCTテストがより適当であったであろう。

また、同実践は、生徒を上位群（2回のプリテストで平均が11点以上のもの）と下位群（11点未満のもの）の2群に分けての結果も出している。下位群の生徒はディベート活動前後（12月と2月）のCTテストの平均点で有意な上昇が認められた（9.21→12.30）が、上位群の生徒は逆に下がっていた（13.00→12.30）。この点に関して、考察では、「英語ディベート活動は、批判的思考をあまりしない生徒の批判的思考を育成するのに効果的な活動と言える。」と結論づけられ、上位群の生徒については、活動の前後で点数が下降したのは「ディベート活動を通してより多視点から問題をとらえ、熟考するようになったためだと考えられる。推論を正確に行うのに以前より時間を必要としたが、回答時間制限のために不正解が増えたのではないか。」との説明が加えられていた。では下位群の生徒は、多視点から問題をとらえたり、熟考したりするようにはならなかったのだろうか、という当然の疑問がわくが、それ以上に、そもそもディベート活動をとおして読解過程における推論の能力が向上するとの考えに妥当性はなく、久原らのCTテストをディベート活動の効果検証に用いたこと自体が間違っていたと考えられる。久原らのCTテストでは同実践による学習者の変容を測定することはできなかったであろう。

5 まとめ

本稿では、久原・井上・波多野（1983）の「批判的思考力テスト」を、英語教育におけるCTの指導実践の効果検証に用いることの妥当性について論じた。久原ら（1983）のCTテストを詳細に分析し、同テストは、批判的思考力を文章の読解過程において働く推論の確かさの微妙な区別をする能力と定義し、「批判的思考力テスト」と称されているが、内容は推論のみに特化した推論のテストであることを明らかにした。（名称だけの独り歩きを防ぐためにも、たとえば少なくとも「批判的思考力（推論）テスト」とでも命名されるべきであった。）さらに、同テストの推論と通常の英語教育における推論とを比較し、久原らのCTテストの推論は、英語教育の読みにおいて使われる、状況モデルの構築を目指した、理解のための推論とはその性質を異にすることも示した。また、CTは複数の構成要素から成る複合的な能力であるため、実践の指導効果の検証をするには、その実践でCTのどの要素に焦点を置いたのかを明示し、その要素を測定できる課題が必要となるが、これまでのいくつかの実践報告では、焦点化したCTの要素が推論であるとは認められないにもかかわらず、久原ら（1983）の推論のテストを効果測定に用いて、その効果を主張している事例もあった。今後、CTの伸長をめざす実践研究においては、実践研究の目的、対象とするCTの構成要素、その要素を測れる課題の3つが整合的に対応していることを確認すべきである。その認識がCTに関わる研究者や実践者に広く共有されることを期待したい。

謝辞

本研究は、JSPS科研費17K03029の助成を受けたものです。

参考文献

- 有嶋宏一・田中瑞穂・鶴田美里映・手島今日子 (2010). 「高校生の英語ディベート活動は英語スピーキング力と批判的思考力を伸ばすのか」 *Step Bulletin*, 22. 財団法人日本英語検定協会. pp. 115-127.
- 伊佐地恒久 (2016). 「テキストを深く読むクリティカル・リーディングの授業」三浦孝・亘理陽一・山本孝次・柳田綾 (編著) 『高校英語授業を知的にしたい—内容理解・表面的会話中心の授業を超えた』研究社.
- 卯城祐司 (編著) (2009). 『英語リーディングの科学—「読めたつもり」の謎を解く』研究社.
- 楠見孝 (2011). 「批判的思考とは—市民リテラシーとジェネリックスキルの獲得」楠見孝・子安憎生・道田泰司 (編) 『批判的思考を育む—学力と社会人基礎力の基盤形成』有斐閣. pp. 8-9.
- 久原恵子・井上尚美・波多野諠余夫 (1983). 「批判的思考力とその測定」『読書科学』第27巻第4号.
- 田中武夫・島田勝正・紺渡弘幸 (編著) (2011). 『推論発問を取り入れた英語リーディング指導—深い読みを促す英語授業』三省堂.
- 平山るみ (2004). 「批判的思考を支える態度および能力測定に関する展望」『京都大学大学院教育学教育研究科紀要』50.
- 道田泰司 (2015). 「批判的思考とはなにか」楠見孝・道田泰司 (編) 『批判的思考—21世紀を生きぬくりテラシーの基盤』新曜社. p. 2.
- 文部科学省 (2011). 学習指導要領「生きる力」https://www.mext.go.jp/a_menu/shotou/new-cs/gengo/1300857.htm
- 文部科学省 (2012a). 「批判的思考について—これらかの教育の方向性の提言—」
https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo3/047/siryo/_icsFiles/afiedfile/2012/09/20/1325670_03.pdf
- 文部科学省 (2012b). 「社会の期待に応える教育改革の推進—教育改革の7つのポイント」
https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/siryo/_icsFiles/afiedfile/2012/06/28/1322862_1.pdf
- Ennis, R., H. (1987). A taxonomy of critical thinking dispositions and abilities. In Baron, J. B., & Sternberg, R. J. (Eds) *Teaching thinking skills: Theory and practice*. W. H. Freeman and Company.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Structuring inferences during narrative text comprehension. *Psychological Review*, 101.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.
- Nuttall, C. (1996). *Teaching reading skills in a foreign language*. Macmillan.
- P21—Partnership for 21st century learning (2009).
http://www.p21.org/index.php?option=com_content&task=view&id=254&Itemid=120

注

- ¹ 新潟県立大学国際地域学部
- ² 新潟医療福祉大学