

# 児童の作文評価をめぐる諸問題

— 観点別到達度評価に関わる史的・測定論的考察 —

梶井芳明

(2002年9月30日受理)

Problems in the evaluation of children's compositions in Japanese schools  
— Historical and theoretical considerations of criterion-referenced evaluation —

Yoshiaki Kajii

This paper discusses the changes of the objectives of the Japanese language education and Japanese writing in the Japanese National Standards after World War II and researches that examined the validity and reliability of the criteria for evaluating writing in the fields of Japanese language education and educational psychology. Results of the review show that objectives of the Japanese language education and Japanese writing have been changing. In the 1947-1968 Japanese National Standards, the objectives were activity-based, that is, developing children's speaking, listening, writing and reading skills. In 1977-1989, however, these objectives have changed to ability-based, focusing on the children's ability to express and understand. And in the New Japanese National Standards (1998), the objectives changed again to activity-based. Likewise, the system for evaluating children's compositions have also changed, that is, the criteria used for each grade level have been integrated, resulting in three sets of criteria; one each for low grade (Grades 1 and 2), middle grade (Grades 3 and 4), and high grade (Grades 5 and 6). Thus, teachers need to use the criteria more carefully, taking into consideration the actual grade level of the children. In terms of research, there is a need to clearly define the evaluation items and for the teachers to experience the use of these items in order to improve the validity and reliability of the evaluation of compositions.

Key words: Japanese National Standards, transition of Japanese writing objectives, criterion-referenced evaluation, validity and reliability of criteria for evaluating writing, use of criteria for evaluation

キーワード：学習指導要領、「書くこと」の目標の変遷、到達度評価、評価項目の妥当性・信頼性、評価観

## 1. はじめに

小学校国語科で行われている作文指導は、学習指導要領に記載された指導内容に沿って行うことが求められており、今日、それらの指導内容を確実に習得したかどうかの評価を一層徹底する必要性が指摘されている（教育課程審議会，2000）。しかし、多くの教員は、評価の基準が曖昧であるが故に、作文評価に対し困難

感を抱いている（国立国語研究所，1978）。これは、学習指導要領に記載された指導内容を反映させた評価項目、いわゆる作文能力の定義が明確でないこと、あるいは定義がなされていたとしても、実践場面では使用しづらいものであることを示唆している。したがって、作文評価研究を進めるに当たっては、学習指導要領の指導内容を反映させた評価項目の妥当性・信頼性の検討を行うことが避けられない課題であるといえる。

そこで本研究では、第1に、我が国における作文評価が、これまでどのような観点のもとで行われてきたのかを明らかにするために、戦後の学習指導要領から国語科の「内容」の領域（以下、内容領域）、および各学年の「書くこと（作文）」の指導目標と観点別指導内容（以下、指導目標・観点別指導内容）の変遷について概観する。第2に、作文評価を困難にする要因を明らかにするために、これまで国語科教育、ならびに教育心理学の分野で行われてきた作文評価研究の中で、評価の妥当性・信頼性を取り扱った研究について概観する。第3に、以上2つの概観の結果をもとに、今後の作文評価研究に求められる課題について考察する。

## 2. 戦後の学習指導要領に見る国語科の内容領域ならびに「書くこと（作文）」の指導目標・観点別指導内容の変遷

### 2.1. 国語科の内容領域の変遷

戦後、学習指導要領に記された国語科の内容領域は、主に「活動」による区分と「能力」による区分がなされてきた。具体的には、昭和22年版（試案）から昭和43年版までは「活動」による区分であったが、昭和52年版、平成元年版では「能力」による区分となった。ところが、平成10年版では、再び「活動」による区分となった。それぞれの時期における領域構成は、次の通りである。

昭和22年版（試案）では、「話すこと（聞くことを含む）」、「つづること（作文）」、「読むこと（文学を含む）」、「書くこと（習字を含む）」、「文法」の5領域で構成されていたが、昭和26年版（試案）では、「聞くこと」、「話すこと」、「読むこと」、「書くこと」の4領域となった。その後、昭和33年、43年版では、「聞くこと・話すこと」、「読むこと」、「書くこと」、「ことばに関する事項」の3領域1事項となった。このように、昭和22年版（試案）から昭和43年版までは「活動」が重視され、さらに各「活動」の目標では、人間形成に関わる習慣、態度、技能、能力を養うことが重視されていた。なお、中沢（1999）は、それぞれの時期の内容領域の区分とその構成から、昭和43年版において「国語の技能とことばに関する事項は、言語活動を通して伸ばされる<sup>1)</sup>」という原則が確立されたことを指摘している。

ところが、昭和52年、平成元年版では、これまでの「活動」による区分が一変して、「表現」、「理解」、「言語事項」の2領域1事項といった、「能力」による区分となった。このような領域区分の変更について、

文部省は、小学校国語科では「内容を整理し、表現力を一層高めるため、現行の3領域（『聞くこと・話すこと』、『読むこと』、および『書くこと』）を『表現』および『理解』の2領域に改め、それに『言語事項（発音・文字に関する事項、文法的な事項など）』を加えて内容を構成する」、「『表現』の領域の内容は、文字言語および音声言語による表現力を養うための基本的な事項を取り上げて構成し、特に文章による表現力を高めることに重点を置く（以下、省略）」と説明している（教育課程審議会、1976）。

昭和43年版では、「『2 内容』（技能とことばに関する事項）は、『3 内容の取り扱い』（言語活動）を方法として学習されるべきものである」といった原則が確立された（中沢、1999）。しかし、昭和52年版、平成元年版では、方法として規定していた「内容の取り扱い」の項が完全に削除された。すなわち、教育課程の基準として規定されていた国語科の「目標・内容・方法」の中から「方法」が取り除かれたのである。

ところが、平成10年版では、再び「話すこと・聞くこと」、「書くこと」、「読むこと」、および「言語事項」の3領域1事項の「活動」を重視した区分となり、「2 内容」の構成も、昭和43年版とほぼ変わらないものとなった。さらに、「2 内容」に記された事項を指導する際の「方法」に相当する「内容の取り扱い」の項も、再び記載されることとなった。なお、構成上著しく異なる点といえば、昭和43年版では、「ことばに関する事項（言語事項）」が各領域ごとにあげられていたのに対し、平成10年版では、「2 内容」の最後に一括してあげられている点である。

以上のことから、これまで国語科の内容領域は、主に「活動」と「能力」といった2つの基準から区分がなされてきたことが確認された。また、「書くこと（作文）」における観点別到達度評価の基準の在り方について検討する上では、「活動」による「目標・内容・方法」の原則が形作られた昭和43年版と、「能力」による区分が完成した平成元年版、さらには現行の平成10年版の指導目標・観点別指導内容の比較が必要不可欠となる。なお、本研究では、昭和52年版は比較の対象として取り上げなかった。それは、昭和52年版と平成元年版は、指導目標・観点別指導内容の記述に大きな違いが見られないからである。そこで本研究では、平成元年版を「能力」による区分が完成した時期として取り扱うこととした。

では、このような国語科の内容領域の変遷のもと、「書くこと（作文）」の指導目標・観点別指導内容は、それぞれどのように変化してきたのであろうか。

2.2. 「書くこと（作文）」の指導目標・観点別指導内容の変遷

昭和22年版の試案から、名称が異なることはあっても、一貫して「書くこと」に関する領域は存在していた。名称は、「作文」（昭和22年版）、「書くこと」（昭和26, 33, 43年版）、「表現」（昭和52, 平成元年版）、「書くこと」（平成10年版）と変化してきた。昭和52年版、ならびに平成元年版では、「表現」といった「能力」を重視した区分となり、国語科の目標（昭和52年版）も、これまでの人間形成を重視した目標から能力形成を重視した目標へと変更した（中沢, 1998）。なお、各領域の提示順序が時代により変化していることについては、それぞれの時期に各領域の重要度が異なることを反映した結果であるという指摘もある（大越, 2002）。

まず初めに、「書くこと（作文）」の指導目標は、戦後、どのように変化してきたのであろうか。昭和43年版、平成元年版、平成10年版のそれぞれの時期における各学年の目標を、大越（2002）を参考に「相手・目的意識」、「題材」、「叙述の仕方」、「技能」、「態度」の5つの観点から整理したのが、TABLE 1である。

TABLE 1. 書くこと（作文）の学年目標の変遷

	昭和43年版	平成元年版	平成10年版
相手・目的意識			相手や目的に応じ(3, 4)/目的や意図に応じ(5, 6)
題材	身近な事柄(1)/身近な事柄(2)	経験したことなど(1)	経験したことや想像したことなど(1, 2)/調べたことなど(3, 4)/考えた事など(5, 6)
叙述の仕方	事柄をまとめて(3)/段落ごとの構成を考えて(4)/文章全体の構成を考えて(5)/文章全体の構成を考えて(6)	(経験したことなどが)分かるように(1)/事柄の順序をはっきりとすることに(2)/表現する内容の要点が分かるように(3)/表現する内容の中心点(4)が分かるように(4)/主題や要旨をはっきりとした表現(5)/目的や意図に応じた表現(6)	順序が分かるように(1, 2)/伝わるように(3, 4)
技能	簡単な文章を書く(1)/順序よく文章を書く(2)/簡単な文章を構成の文柄をまとめる(3)/事柄のある文章を書く(4)/主眼点や要旨をはっきりとした文章にふさわしい文章を書く(6)	文と文をつなげたりする(1)/文の続き方を意識して文章を書く(2)/事柄のまとまりを構成する(3)/段落を構成する(4)/意図を明らかにしたり(5)/全体を構成し(6)	文の書き方に注意して文章や段落を相互の関係などを工夫して書く(3, 4)/筋道を立てて文章を書く(5, 6)
態度		進んで表現しようとする(1)/正しく表現しようとする(2)/分かりやすく表現しようとする(3)/内容を整理して表現しようとする(4)/相手や場の状況を考慮して表現しようとする(5)/適切で効果的な表現しようとする(6)	楽しんで表現しようとする(1, 2)/適切に表現しようとする(3, 4)/効果的に表現しようとする(5, 6)

注1：( )内の数字は、学年を示す。  
注2：観点点、大越(2002)を参考に選定した。

TABLE 1より、観点の数は、改訂の度に増加していることが分かる。具体的には、平成元年版からは「態度」に関する観点、平成10年版からは「相手・目的意識」に関する観点が新たに盛り込まれている。また、各観点の内容は、平成元年版が最も詳しい記述である。なお、平成10年版では、各学年の指導目標が2学年ずつのまとまりとなり、さらにそれぞれの観点の内容は、上級学年（第2, 4, 6学年）に合わせたものになっている。

次に、「書くこと（作文）」の観点別指導内容を年代別に示したのがTABLE 2である。TABLE 2より、観点別指導内容を年代別に比較すると、昭和43年版では、「興味・関心・意欲」、「表現語句」、「辞書の利用」、平成元年版では、「関連指導」、「視写・聴写」といった観点からも分類、記述されているのが特徴的である。また、観点の数は、昭和43年版で9つ、平成元年版で8つ、平成10年版では5つとなっている。なお、平成10年版の指導内容は、指導目標（TABLE 1）の場合と同様、それぞれ2学年ずつのまとまりになっている。

以上のことから、平成10年版では、「書くこと（作文）」の指導目標・観点別指導内容が、これまでの第2, 4, 6学年の指導目標・観点別指導内容に到達度基準を合わせた2学年ずつのまとまりとなり、さらに、観点別指導内容の観点が減少していることから、従来にも増して、観点を明確にして、各学年ごとに発達に応じた弾力的な指導・評価を行うことが実践現場に求められているといえる。なお、「書くこと（作文）」の活動を通してその能力を高めていくには、作文の能力構造、いわゆる観点別到達度評価の基準について検討を行うことが重要となる。このことから、平成元年版をもとに作文の能力構造を明らかにし、さらに昭和43年版を参考にしながら、能力の向上が保証された活動の在り方について検討していくことが重要な課題といえよう。

2.3. 観点別到達度評価とは

日本において「到達度評価」という用語が初めて文献に登場したのは、1975年に京都府教育委員会が作成した「研究討議のための資料・到達度評価への改善を進めるために」である（望月, 2001）。そこでは、学習指導要領の中から、子どもに習得させるべき基本的指導内容（評価観点・項目）を、全体指導計画、年間指導計画、および授業展開計画のそれぞれのレベルに応じて精選、設定し、それを授業過程の形成的評価、単元終了時の総括的評価、さらには学年末の総括的評価の各時期で実施した後に、そこで得られた情報を次なる計画に役立てていく（「京都の国語教育・到達度評価の実践」編集委員会, 1978；水川隆夫, 1981）。



このように、到達度評価の評価観点・項目の具体化、評価実践による有効性の検証、得られた情報のフィードバックといった循環的な教育活動を通して、評価観点・項目のみならず、評価活動そのものの質が向上するといった道筋が提唱された。このような、到達度評価の理念に基づき、客観的・科学的な児童・生徒の基礎学力の充実に資し得る教育体系を、高木(1981)は、「到達度評価教育システム」と呼んでいる。

以上のような、評価観点・項目の具体化を伴う到達度評価、いわゆる観点別到達度評価が、教育実践現場に正式に導入され始めたのは、文部省(1980)が、各都道府県教育委員会あてに示した「小学校児童指導要録及び中学校生徒指導要録の改訂について(通知)」からである。この中で、「各教科の学習の記録」の部分が大幅に改められた。中でも、「評定」欄は従来とほぼ同じ考え方で設定されているものの、新たに「観点別学習状況」欄が設けられて、目標に基づいて学習到達度を評価するようになった。

井上(1979, 1981)は、小学校教員を対象に、教育評価に対する意味構造を意味微分法により分析した。その結果、教員は、年代によらず、相対評価に比べ観点別到達度評価に対してより好意的なイメージをもつ傾向が強いことが示唆された。さらに、観点別到達度評価については、理念としてはよいものであると評価しているが、実際の指導や評価に用いるとなると、大変な困難を伴うと認識していることが示唆された(井上, 1993)。

これらのことから、観点別到達度評価に対する困難観を克服し、「到達度評価教育システム」(高木, 1981)を有効に機能させるためには、評価観点・項目の具体化と、その有効性の検証作業を、授業過程、単元終了時、さらには学年末といった各時期で体系的に行う必要があるといえる。

では、評価観点・項目の具体化と、その有効性の検証作業は、これまでどのように行われてきたのであろうか。

### 3. 国語科教育の分野における作文評価研究

これまで、国語科教育の分野で行われてきた作文評価研究においては、作文の評価結果には評定者個人の作文に対する好みといったバイアスが加わりやすかったことが、実証的研究により明らかにされてきた。

例えば、有定(1969)は、7名の教員を対象に、小学校5年生の子どもが書いた生活文を10の評価項目に沿って評定させる課題を行った。その結果、作文の内

容面に関わる項目(ものの見方・感じ方・考え方、発見・創造、個性的表現、統一性・集中性、追求性、具象性)では、評定者間で評定結果が一致しにくいのに対し、作文の形式面に関わる項目(段落、接続語・指示語の使用、漢字の使用、文字量)では、評定者間で評定結果が一致しやすいことが示された。

また、大熊(1996)は、教員養成系大学の国語専修の学生と、専門研究分野を国語とする小学校教員を対象に、小学校1年生から4年生までの子どもが書いた6編の詩の中から、最もよいと思われるもの1編を選出させた。学生と教員の結果をもとに、それぞれの詩の選出回数順位を比較したところ、学生は、詩らしい表現テクニックがある、いわゆる上手な詩を高く評価するのに対し、教員は、五官(原文ママ)を十分に働かせて、飾らずにありのままに一生懸命書いた詩、ほほえましい詩、子どもにしか書けない詩を高く評価する傾向が強いことが示唆された。

これらの研究から、作文評価においては、評価者の主観的な判断によらず、ある程度の妥当性を保って評価を行うこと、いわゆる客観的な評価を行うことが非常に困難であると推察される。では、評価者の主観的判断をできる限り排除し、作文評価を客観的で妥当なものにする試み、すなわち評価観点・項目の具体化と、その有効性の検証作業は、どのように行われてきたのであろうか。

これまで、国語科教育の分野では、主に評価観点・項目の具体化と、それをういた作文指導実践が数多く行われてきた(e.g., 江口, 1972; 平井, 1971; 倉沢, 1952; 森岡, 1963)。ところが、それら評価観点・項目の有効性を実践場面を通して検証した研究は数少ない。そこでの検証は、ある作文について、複数の評価者に独自に設定した観点に基づいて評語を付けさせ、評語内容について評定者間での一致度を見る観点の信頼性の検討(e.g., 川上・斎藤・大矢, 1968; 森岡, 1968)や、評語内容から、どのような観点に基づいた記述がなされているのかを見る評語の妥当性を検討したもの(e.g., 市毛・須田, 1988; 倉沢, 1952, 1987)がほとんどで、具体化した評価観点・項目そのものの妥当性・信頼性を数量的に検討した研究は見当たらない。これは、数量的な研究手法が国語科教育に普及していなかったことに加え、日本の作文教育が、従来、言語技術といった技能的側面よりも、子どもの人格の形成といった情意的側面を重視する傾向が強かったことに関係していると思われる(井上, 1984)。これに対し、アメリカでは、B. S. Bloomの「教育目標の分類学(タクソノミー)」に基づき、作文能力の構造分析に関する研究が数多く行われてきた(Foley, 1971/

1974)。このような文化的な背景を差し引いて考えてみても、我が国の国語科教育における作文能力の構造的な分析、すなわち評価観点・項目の妥当性・信頼性の検討に関する研究の数は、国語科教育全体の研究数から見て相対的に少ないのが実状のようである。

## 4. 教育心理学の分野における作文評価研究

### 4.1. 作文能力を測定・評価する試み—間接的な評価と直接的な評価—

これまで、教育心理学の分野で行われてきた作文評価研究においては、作文能力（以下、作文力）を数量的（客観的）に測定・評価する方法の開発が、実践場面を通して数多く行われてきた。具体的には、作文力の評価について、間接的な評価（以下、間接的評価）と直接的な評価（以下、直接的評価）とを比較する研究が行われてきた。間接的評価では、作文力が間接的な尺度、いわゆる標準化客観テスト（standardized objective tests）（以下、客観的テスト）により評価され、直接的評価では、作文力がある課題の下で書かせた作文に対する総合評価により評価された。なお、岡本（1966）は、これら2つの作文力の測定・評価方法の長所と短所を、次のように整理している。「客観的テストによる方法は、信頼度（原文ママ）は高く、測定に一貫性がある。しかし、測定できる作文力の面が片寄っていて、妥当性については一般に難点がある。つまり、客観的テストによる方法では、句読法、慣用語、文法などに関する能力はよく測定できるものの、素材内容や文章全体の構成面の能力は、十分に測定することができない。それに対し、ある文題について実際に書いた作文を用いる方法では、作文力を多面的に測定することができる。しかし、この方法においても、作文の実施法や採点法について十分に研究し、種々の測定条件を明確に規定しておかないと、信頼度や妥当性を高めることができない。」（岡本、1966）。

以下、作文力の間接的評価と直接的評価とを比較した研究を紹介する。

Breland & Gaynor（1979）は、通年の英作文の授業を受講する学部1年生の学生（2000名以上）を対象に、前期の最初と最後、および後期の最後の授業時に、作文力の間接的評価と直接的評価を実施し、2つの評価の相関が高いことを示した。具体的には、間接的評価は、TSWE（Test of Standard Written English）を用いた課題、直接的評価は、3種類の説得的文章（persuasive writing）をそれぞれ20分の制限時間の下で書かせる課題であった。直接的評価を意図した作

文の評価は、15名の大学の英語教授と高等学校の英語教師が行い、それぞれの評定者は、各作文について最低2人で個別に評定を行うよう任意に割り当てられた。なお、評定は、6件法による総合評定で行われ、評定に際しては、事前に練習課題が設けられた。その結果、全評定者の3種類の文章についての間接的評価と直接的評価を、それぞれ合計した評定平均値をもとに、時期を込みにして2つの評価の相関係数を求めたところ、.76と強い相関が見られた。また、各時期ごとに、2つの評価の相関係数を求めたところ、.52から.76といった、中程度から強い相関が見られた。

Hogan & Mishler（1980）は、MAT（Metropolitan Achievement Tests）の研究指定校となった4つの学校の3年生（9歳）と8年生（14歳）の児童・生徒を対象に、作文力の間接的評価と直接的評価の相関を求めた。その結果、小学校段階においても間接的評価と直接的評価の相関が高いことが示され、Breland & Gaynor（1979）の大学生を対象に行った調査と同様の結果が得られた。

Moss, Cole & Khampalikit（1982）は、オハイオ州の一地方学区の4年生（10歳）と7年生（13歳）、および10年生（16歳）の児童・生徒を対象に、作文力の間接的評価と直接的評価の相関を求めた。その結果、10年生（高校時）に比べ4、7年生（小学時から中学時）の児童・生徒では、間接的評価と直接的評価の相関が低いことが示され、Hogan & Mishler（1980）の結果と異なる結果が得られた。また、分析的評価と総合評価の結果をもとに、児童・生徒×課題×評定者の分散分析を行い、各学年ごとにそれぞれの要因の分散成分の推定値（variance component estimate）を求めたところ、それぞれの評価で、10年生を除いて、誤差要因の次に児童・生徒×課題の要因の値が最も大きいことが示された。一方、10年生では、誤差要因の次に児童・生徒の要因の値が最も大きいことが示された。

Boodoo & Garlinghouse（1983）は、34名の大学生を対象に、肯定的または否定的な立場に立ち、授業で習った3つの心理学理論についての説得的文章（persuasive writing）を書かせる課題を行った。調査には、教育心理学入門の授業、計7コマが使用された。それぞれの課題内容は、前の週に行われた授業内容、および授業終了時に行われた客観的テストの内容と同じものであった。評定者は、これら一連の調査内容を熟知した3名の教師で、それぞれの作文に対し5件法による総合評定を行った。その結果、学生×課題×評定者の交互作用、すなわち誤差要因の値が最も大きいことが示され、以下、学生×課題の要因、学生の要因、評定者の要因の順であった。さらに、各課題の

作文について、それぞれの評定者が与えた評定値をもとに因子分析を行ったところ、課題1と2の間には正の相関が見られたが、課題3と課題1、2の間には、それぞれ負の相関が見られた。また、客観的テストの結果においても同様に、課題間で正答率に違いが見られた。一方、評定者を込みにして、各課題の作文に与えられた評定値合計をもとに課題間での相関係数を求めたところ、それぞれの課題間では.20から.30の正の相関が見られた。これらのことから、評定者を込みにして課題間の相関を求めることが、課題間の関連についての解釈を歪める危険性があることが示唆された。

以上の作文力の間接的評価と直接的評価とを比較した研究の結果から、以下の2点が示唆された。

第1に、テストの成績がよいからといって、必ずしも作文を上手に書けるとは限らないことである。これは、間接的評価と直接的評価とでは、共通に評価できる内容に制限があることを意味している。

例えば、Harris (1963/1966) は、作文の指導方法の違いにより、指導後の間接的評価と直接的評価の成績が異なることを明らかにした。具体的には、中学生を対象に、文法教科書による正規の文法時間を設けての指導と、作文の時間内での書かれた作文についての機能的文法指導<sup>2)</sup>とは、どちらが作文力の向上に役立つのかを検討した。その結果、1つに、中学校の段階では、正規の文法時間を設けて伝統的な文法指導を行うことは、文法テストの成績(間接的評価)の向上には役立つが、直接、作文力(直接的評価)を向上させることにはあまり役立たないこと、2つに、直接、作文力を向上させるためには、作文の時間に、作文に即して機能的に文法を指導する方が有効であることが示唆された。

指導方法が異なるということは、方法を支える指導目標も異なることを意味する。Harris (1963/1966) の調査からは、指導目標に対応した評価の在り方、すなわち指導と評価の一体化の重要性を支持する結果が得られたと解釈できよう。

第2に、課題に対する習熟度といった書き手の要因は、評定結果に大きな影響を及ぼす傾向が強いことである。このことから、作文の評価に当たっては、客観的テストの項目の分析と同様に、課題の内容についても児童・生徒(学生)の実態に即して分析すること、すなわち、あらかじめ作文課題で育成される能力構造を、書き手の実態に即し体系的に定義しておくことが、評価の妥当性を向上させる上で重要であるといえる。

#### 4.2. 作文の評定構造—評定点に影響を及ぼす書き手の能力特性—

まず初めに、評定者は、作文に表れる書き手の能力

特性をどのようにとらえるのか、また、その中でも、主にどのような特性を反映させて評定を行うのかについて分析をした研究を紹介する。

Remondino (1959) は、作文の評定点に関与する複雑な要因を因子分析を用いて整理し、作文の評定構造を研究した。作文の評定経験に個人差がある10名の評定者(経験あり:  $N=4$ , 経験なし:  $N=6$ )それぞれに、職業専門学校に在籍する15、16歳の学生が自分自身のことについて書いた作文、計230編を、17の観点から評定させ、評定者ごとに因子分析を行った。分析結果から、経験の有無によらず、評定構造は4つの因子(外見上の美しさ、用語力、内容と構成、内容の個人的側面)でほとんどが説明できると結論している。

Remondino (1959) の研究をもとに、渡部・平・井上(1988) は、2種類の作文のデータを用いて、それぞれ15、17観定の構造を、複数の評定者がプールされた相関行列から分析した。その結果、一方のデータ(15観点)では2つの因子(読後感想的な評価に関する観定点群、文法や用語力などの分析的評価に関する観定点群)、もう一方のデータ(17観点)では3つの因子(読後感想に関する観定点群、文法・用語に関する観定点群、文の構成に関する観定点群)に分類されたとしている。

次に、分析的評価が総合評価へ及ぼす影響を検討した研究を取り上げる。これらの研究は、作文に与えられる評定点が、書き手のどのような能力と関係しているのかを考える上で重要であると思われる。

Harris (1977) は、36名の高等学校の英語教師を対象に、12名の学生の作文を添削させる課題と、12編の作文を順位づけ(rank order)させた後に、質問紙法により、順位づけの際に重視した観点をあげさせる課題を実施した。添削課題の結果からは、教師は作文評価の際に、技巧を重視する傾向が強いことが確認され、一方、質問紙法による解答結果からは、内容と構成を重視する傾向が見られた。これらのことから、作文評価の実践と理論との間では、重視される評価観点に矛盾が生じやすいことが示唆された。

Freedman (1979, 1992a, 1992b) は、学生が書いた96編の評論文を、あらかじめ内容、構成、文構造、技巧の4観点から、上手または下手の2つの基準で書き直しを行い、論文の質について英語教育に熟達した12名の教師に総合評定(4件法)させた。総合評定の後には、それぞれの作文について、観点別に上手か下手かを判断させる課題を行った。総合評定の結果をもとに、内容×構成×文構造×技巧の分散分析を行い、各要因の分散成分の推定値を求めたところ、内容の要因が最も大きく、かつ主効果が有意であった。その次に大きかったのが、構成(主効果が有意)で、技巧は最

も小さく主効果に有意差は見られなかった。また、構成と技巧、および構成と文構造の交互作用がそれぞれ有意であった。これらのことから、評論文の評定結果に大きな影響を及ぼす要因は、内容と構成であること、また、技巧と文構造は、構成が上手に書かれている時に、それぞれ妥当な評定が行われる傾向が強いことが示唆された。さらに、それぞれの作文に対する書き直しの基準（上手、下手）と、総合評定後に行われた観点別の判断（上手、下手）とを比較したところ、内容、構成、文構造、技巧の順に、書き直しの基準と観点別の判断とが一致しにくい傾向が示された。このことから、文構造や技巧に比べ、内容や構成は、評定者の主観的な判断が入りやすい観点であることが示唆された。

渡部他（1988）では、総合評価と最も関係が強いのは読後感を評価する項目であり、最も関係が弱いのは誤字・脱字の有無を評価する項目であることが示された。また、評定者ごとに総合評価を基準変数、分析的評価を説明変数として重回帰分析を行ったところ、重相関係数、および調整済み重相関係数の値は、それぞれ.22から.64、.09から.61と低いことが示された。

石田・森（1985）は、大学生と小学校教員との間で児童の作文（生活文）の評価方法に違いが見られることを明らかにした。それによると、大学生が、客観的には構成力（文章全体の構成のうまさ）を重視し、主観的には説得力（児童の気持ちがよく伝わるか）を重視する傾向が強いのに対し、教員は、客観的には説得力を重視し、主観的には表現力（個々の文の表現や言いまわしのうまさ、巧みさ）を重視する傾向が強いことを指摘している。

以上の諸研究から得られた知見を、一義的に解釈するのは難しい。なぜならば、それぞれの作文に与えられた評定点は、あらかじめ設定した評価観点・項目を十分に反映しているとは言い難いと推察されるためである。また、総合評価として得られた評定点を分析的に説明しようとしても、十分には説明しきれない部分が多いことも示唆された。

#### 4.3. 作文の評定の信頼性—評定者の要因—

平・江上（1992）は、作文評価における評定者要因に関わる問題を、以下の2点から整理している。

- (1) 評定者が、どのような評価観点・項目から作文を評定しているのかについて。
- (2) 評定者が、評定の信頼性にどのような影響を及ぼしているのかについて。

(1)は、評定結果が、書き手のどのような能力を反映したものなのかについてで、平・江上（1992）は、総合評価は分析的評価に比べ評定者に対する負担は軽い、そこから得られる情報はその性質上、評定の構造

を分析する手だてがないことから、分析的評価に対するアプローチが、総合評価の構造解析という点からも重要であるとしている。(2)は、評定者の主観と評定の信頼性の関係についてで、平・江上（1992）は、同一の評定者がある1つの作文を繰り返し評定する際の信頼性（再評価信頼性）と、異なる評定者がある1つの作文を評定する際の信頼性とに分けて考える必要性を指摘している。

渡部他（1988）は、現役教師、および大学院生を対象に、2つの作文評定課題を実施し、同一の作文（小論文）（データⅠ：165編、データⅡ：124編）を8人の評定者（データⅠ：N=3、データⅡ：N=5）それぞれに、2回ずつ評定させた。2回の評定平均値の差の検定を用いて再評価信頼性を調べたところ、8人中5人の評定者が有意であった。同一の評定者による繰り返しの相関係数を求めたところ、.04から.91の間に広く散らばっていた。また、評定者を2人ずつ対にして評定値間の相関係数を求めたところ、.05から.58であった。さらに、一般化可能性理論を応用して、一要因配置モデルの下で、 $\rho = .90$ の信頼性を得るために必要な評定者の人数を求めたところ、35人（データⅠ：N=11）と231人（データⅡ：N=7）であった。これらのことから、8人の評定者はそれぞれ異なった基準により評定を行ったこと、また、その基準は非常に不安定なものであったことが示唆された。これは、作文のよさは一義的に定まるものではなく、個人内、および個人間で不安定なものであることを意味している。なお、十分な信頼性を確保するためには、実際の作文評価場面とはかけ離れた膨大な人数の評定者を確保しなければならないことも示唆された。

それでは、評定者の負担をそれほど増やすことなく、評価の信頼性を高める方法はないのであろうか。

石井（1981）は、小論文入試の信頼性を高めるためには、作文の評定に先立って、①出題の意図と評価の基準を明確にすること、②評定者が訓練・経験を積み重ねること、が必要であることを指摘している。

また、平（1995）では、合計167名の中学生と高校生が書いた物語文を、現役高校国語教師5名と元高校国語教師1名、および書籍編集者1名の、計7名に評定させたところ、分析的評価と総合評価における評定の信頼性係数が、ともに.84と高いことが示された。このような高い信頼性係数が得られた原因について平（1995）は、1つに、課題に合わせた分析的評価項目を設定することにより、評定の仕方を明確に規定できたこと、2つに、エキスパートの評定者を確保できたことによると分析している。

これらの研究から、作文評価の妥当性・信頼性を高

めるためには、①評価の基準(項目)を明確にすること、②評価の経験の積み重ねを多くもつこと、の2点が重要であることが示唆された。すなわち、作文評価においては、指導目標に合致し、作文の内容、および構成的的確さを区別でき、どの部分を改善すべきであるかの指標(情報)を子どもにフィードバックできる、系統的かつ具体的な評価観点・項目が必要となる。さらに、それらの評価観点・項目をもとに、評価の経験を多く積み重ねていくことが、評価の妥当性・信頼性を向上させる上での第一歩となるといえよう。

## 5. 総合的考察

本研究では、戦後の学習指導要領(国語科)の内容領域、および「書くこと(作文)」の指導目標・観点別指導内容の変遷と、国語科教育、ならびに教育心理学の分野で行われてきた作文評価研究の中で、評価の妥当性・信頼性を取り扱った研究について概観した。

その結果、平成10年版の学習指導要領に基づく作文指導と観点別到達度評価を行う上では、従来にも増して、「到達度評価教育システム」(高木, 1981)といった指導計画立案と授業実践の循環的な過程の中で、子どもの実態(発達)に即して「書くこと(作文)」の能力構造を明確にし体系化していくこと、すなわち学習指導要領を反映させた評価項目の妥当性・信頼性を検討していくことが重要であることが示唆された。

例えば、Chase(1986)は、実際の作文評価場面においては、さまざまな要因が複雑に影響を及ぼし合っていることが予想されることから、各要因の主効果に加え、要因間の交互作用の結果を蓄積していくことが作文評価のメカニズムを明らかにする上で重要であると指摘している。

そこで今後、教育心理学の分野での作文評価研究に求められる課題の1つとしては、学習指導要領を反映させた評価項目に沿って作文を評定させる課題を実施し、評定者の要因や作文の質的要因、さらには文種といった多様な要因計画の下で、各要因間の交互作用の結果を蓄積していくのと同時に、評価項目の妥当性・信頼性を検討していくことがあげられる。なお、以上のような要因計画の下で得られた調査結果を、作文力の構造(マトリックス)についての基礎的資料の作成や、作文評価を行う上での留意(注意)事項の提案に役立てていくことも、今後の重要な課題の1つとして残されている。

## 【引用文献】

- 有定稔雄 1969 書くことの評価と留意点 教育科学 国語教育, No.134, 19-25.
- Boodoo, G. M., & Garlinghouse, P. 1979 Use of the essay examination to investigate the writing skills of undergraduate education majors. *Educational and Psychological Measurement*, **43**, 1005-1014.
- Breland, H. M., & Gaynor, J. L. 1979 A comparison of direct and indirect assessments of writing skill. *Journal of Educational Measurement*, **16**, 120-128.
- Chase, C. I. 1986 Essay test scoring: Interaction of relevant variables. *Journal of Educational Measurement*, **23**, 33-41.
- 江口季好 1972 綴方の鑑賞と批評—作品の見かた・考えかた— 百合出版
- Foley, J. J. 1971 Evaluation of learning in writing. In B. S. Bloom, J. T. Hastings, & G. F. Madaus (Eds.), *Handbook on formative and summative evaluation of student learning* (pp.767-813). New York: McGraw-Hill. 石原敏道(訳) 1974 作文教育における学習の評価 渋谷憲一・藤田恵璽・梶田叡一(監訳) 教育評価ハンドブック(上) 第一法規 Pp.259-322.
- Freedman, S. W. 1979 How characteristics of student essays influence teachers' evaluations. *Journal of Educational Psychology*, **71**, 328-338.
- Freedman, S. W. 1992a How characteristics of student essays influence teachers' evaluations. In R. Hayes (Ed.), *Reading empirical research studies: The rhetoric of research* (pp.322-340). Hillsdale, NJ: Erlbaum.
- Freedman, S. W. 1992b Why do teachers give the grades they do? In R. Hayes (Ed.), *Reading empirical research studies: The rhetoric of research* (pp.341-347). Hillsdale, NJ: Erlbaum.
- 藤原 宏(編著) 1990 新学習指導要領「国語」指導内容系統表—小学校・中学校・高等学校— 明治図書 Pp.10-15.
- Harris, R. J. 1963 Summaries of selected research. In R. Braddock, R. Lloyd-Jones, & L. Shoer (Eds.), *Research in written composition* (pp.70-83). Champaign, IL: National Council of Teachers of English. 内藤善康(訳) 1966 アメリカにおける作文指導の研究児童心理, No.235, 167-172.
- Harris, W. 1977 Teacher response to student writing: A study of the response patterns of high school English teachers to determine the basis for teacher judgment of student writing. *Research in the Teaching of English*, **11**, 175-185.
- 平井昌夫 1971 文章評価法 至文堂

- Hogan, T. P., & Mishler, C. 1980 Relationship between essay tests and objective tests of language skills for elementary school students. *Journal of Educational Measurement*, **17**, 219-227.
- 市毛勝雄・須田 実(編著) 1988 「評語」で変わる子どもの作文 小学4-6年編 明治図書
- 井上正明 1979 教師の教育評価に対する意識構造 (I)-SD法による年代差の分析- 福岡教育大学紀要, **29**, 105-121.
- 井上正明 1981 教師の教育評価に対する意識構造 (II)-管理職教師群と一般教師群との比較- 福岡教育大学紀要, **31**, 133-148.
- 井上正明 1993 新しい学力観に基づく評価論と教師の力量 福岡教育大学紀要, **42**, 243-258.
- 井上尚美 1984 国語科における評価の問題 井上尚美・田近洵一・根本正義(編著) 東京学芸大学公開講座III 国語科の評価研究教育出版 Pp.8-17.
- 石田 潤・森 敏昭 1985 小学生の文章表現の発達的变化 広島大学教育学部紀要, **33**, 125-131.
- 石井 巖 1981 「論文試験とその評価」について行動計量学, **8**, 22-29.
- 川上 繁・斎藤喜門・大矢武師 1968 作文評価事例集 小学校 中学校 高等学校 森岡健二・永野賢・宮地 裕・市川 孝(編著) 作文講座5 文章の評価 明治書院 Pp.26-210.
- 国立国語研究所 1978 国立国語研究所報告67 児童の表現力と作文 東京書籍
- 倉沢栄吉 1952 作文教育の体系 金子書房
- 倉沢栄吉 1987 作文の教師-指導法の手引き- 国土社
- 教育課程審議会 1976 小・中・高等学校の教育課程の基準の改善について(答申) 文部省
- 教育課程審議会 2000 児童生徒の学習と教育課程の実施状況の評価の在り方について(答申) 文部省「京都の国語教育・到達度評価の実践」編集委員会(編) 1978 京都の国語教育・到達度評価の実践 地歴社
- 水川隆夫 1981 国語科到達度評価の理論と方法 明治図書
- 望月謙二 2001 到達度評価 日本国語教育学会(編) 国語教育辞典 朝倉書店 Pp.284.
- 文部省 1980 小学校児童指導要録及び中学校生徒指導要録の改訂について(通知)
- 森岡健二 1963 文章構成法-文章の診断と治療- 至文堂
- 森岡健二 1968 文章の評価と作文の評価 森岡健二・永野賢・宮地 裕・市川 孝(編著) 作文講座5 文章の評価 明治書院 Pp.1-25.
- Moss, P. A., Cole, N. S., & Khampalikit, C. 1982 A comparison of procedures to assess written language skills at grades 4, 7, and 10. *Journal of Educational Measurement*, **19**, 37-47.
- 中沢政雄 1998 二十年後戻りした新学習指導要領の国語科教育(1)-見通しを誤っていた文部省- 国語情報, **29**(12), 1-18.
- 中沢政雄 1999 二十年後戻りした新学習指導要領の国語科教育(3)-見通しを誤っていた文部省- 国語情報, **30**(2), 1-19.
- 岡本奎六 1966 作文力の客観的測定 児童心理, No.235, 159-166.
- 大越和孝 2002 小学校学習指導要領の国語科の目標の変遷II 書くことの学年目標 東京家政大学研究紀要, **42**, 63-71.
- 大熊 徹 1996 教育としての評価の実践上の課題-作文教育に関する考察- 田近洵一(編) 国語科の新視点⑥ 自ら学ぶ力を育てる国語科の評価 国土社 Pp.200-212.
- Remondino, C. 1959 A factorial analysis of the evaluation of scholastic compositions in the mother tongue. *British Journal of Educational Psychology*, **29**, 242-251.
- 平 直樹 1995 物語作成課題に基づく作文能力評価の分析 教育心理学研究, **43**, 134-144.
- 平 直樹・江上由美子 1992 ESSAY TEST の方法的諸問題に関する研究の動向について 教育心理学研究, **40**, 108-117.
- 高木一郎 1981 到達度評価とその生かし方 図書文化
- 渡部 洋・平由美子・井上俊哉 1988 小論文評価データの解析 東京大学教育学部紀要, **28**, 143-164.

## 【注】

- 1) 国語の技能, ことばに関する事項, 言語活動は, それぞれ「聞くこと・話すこと」, 「読むこと」, 「書くこと」の各領域の「2 内容」の「(1)次の事項について指導する」以下に記された「技能」, 「(2)(1)の指導と関連させて, ことばに関する次の事項について指導する。」以下に記された「ことばに関する事項」, 「3 内容の取り扱い」以下に記された「言語活動」を示している。
- 2) 機能的文法指導では, 教師は, 生徒の書いた作文の文章構造や文の組み立てをよく検討し, 学級に共通な誤りを発見したならば, 実例と模倣による学習法によって練習を行わせた。その際, 形式文法の立場からの抽象化や一般化を行う学習指導は, 一切行わなかった。

付記 本論文をまとめるに当たり御指導下さいました広島大学森敏昭教授, 森田信義教授, 英語文献の読解の際に有益なアドバイスを下さいました広島大学大学院教育学研究科 R. S. Gravoso 氏に心から感謝申し上げます。

(主任指導教官 森 敏昭)