

小学校高学年児童の文字学習能力と その自己評価について： リッカート尺度と Can-Do 尺度による自己評価の比較

A Study on the Relationship between Letter Knowledge and its
Self-Evaluation among Sixth Graders in Public Elementary Schools:
Comparing Likert Scale and Can-Do Scale Self-Evaluation

アレン玉井光江
Mitsue ALLEN-TAMAI

1. はじめに

新学習指導要領の導入により、2020年度より公立小学校では高学年には外国語、また中学年には外国語活動が導入された。外国語は教科であり、各自治体で文部科学省検定の教科書が採択され、新しい観点に基づく評価も実施されている。一方外国語活動は領域扱いとなり、文部科学省作成の副読本が使われており、文章の記述による評価が行われている。公立小学校で外国語が始まった2020年は、コロナウイルス感染拡大防止のため春季休業が明けるとまでの一斉臨時休業、続いて緊急事態宣言が発出され教育関係者は多くの問題に直面し、新教科への対応どころではなかった。小学校英語は導入3年目を迎えているが、いまだに多くの課題を抱えながら、試行錯誤が続いている。

そのような中、本稿では教科として導入された外国語の指導の中で難しいとされている「読むこと」「書くこと」について焦点をあて、その目標と評価について考えてみたい。特に学習者中心の学習方法や自己調整学習などが注目を集める現在、授業中に簡単に実施できる自己評価については大きな関心が寄せられている。そこで本研究では児童の文字に関する知識やスキルを測定し、彼らのそれらに対する自己評価と比較し、自己評価の信頼性または正確性について2つの研究から検証している。具体的には研究1では能力測定とともに1つのグループではリッカート尺度による自己評価、もう一つのグループではCan-Do尺度による自己評価を実施し、その正確性を比較した。研究2ではCan-Do尺度での自己評価の信頼性を再び検証するとともに、5年生と6年生を対象にその学年差、つまり学習歴の違いに

より自己評価にどのような違いがでるのかを検証した。

2. 文献研究

ここでは新しく教科となった公立小学校での外国語科における「読むこと」「書くこと」の目標を見たのち、その評価について文部科学省から出された資料を基に考えていく。また本論文の主要なテーマである自己評価、また Can-Do 評価についてみていきたい。

2.1 「読むこと」「書くこと」の指導とその目標について

ここでは文部科学省が示している小学校の外国語の「読むこと」「書くこと」についての目標を見たのち、英語圏におけるリーディング能力に関する研究を概観し、日本人の小学生が獲得すべき「読む」力について論考する。2017年告知の学習指導要領では、小学校の外国語での「読むこと」「書くこと」について下記のような目標が設定されている。

【読むこと】の目標

- ア 活字体で書かれた文字を識別し、その読み方を発音することができるようにする。
- イ 音声で十分に慣れ親しんだ簡単な語句や基本的な表現の意味が分かるようにする。

【書くこと】の目標

- ア 大文字、小文字を活字体で書くことができるようにする。また、語順を意識しながら音声で十分に慣れ親しんだ簡単な語句や基本的な表現を書き写すことができるようにする。
- イ 自分のことや身近で簡単な事柄について、例文を参考に、音声で十分に慣れ親しんだ簡単な語句や基本的な表現を用いて書くことができるようにする。

まずは「読むこと」の目標であるが、アルファベットの文字を字形として認識し、さらにその読み方（これ以降名称とする）を理解し、発音できることが最初の目標となっている。名称については「この目標における「読み方」とは音ではなく、文字の名称の読み方を指していることに留意する必要がある。」（文部科学省、2018, p. 78）と、明らかに文字の名称読み（/ei/, /bi:/ など）と音読みを区別し、名称読みのみが（ア）の目標である。

次の（イ）の目標については「中学年の外国語活動の「聞くこと」、「話すこと」の学習活動を通して、音声で十分に慣れ親しんだ簡単な語句や基本的な表現の意味が分かるようにすることを示している。日常生活に関する身近で簡単な事柄について、掲示、パンフレットなどから自分が必要とする情報を得たり、絵本などに書かれている簡単な語句や基本的な表現を識別したりするなど、言語外情報を伴って示された語句や表現を推測して読むようにすることを示している。」（文部科学省、2018, p. 78）と説明されている。この目標については、

「英語の文字には、名称と音がある。児童が語句や表現の意味が分かるようになるためには、当然のことながらその語句や表現を発音する必要があり、文字の音の読み方は、そのための手掛かりとなる。したがって、ここで示されている目標に関して指導する際には、児童の学習の段階に応じて、語の中で用いられる場合の文字が示す音の読み方を指導することとする。その際、中学校で発音と綴りを関連付けて指導することに留意し、小学校では音声と文字との関連付ける指導に留めることに留意する必要がある。」（文部科学省，2018，p. 78

下線は筆者による）と、説明が続いている。つまり、アルファベットの文字に関しては、その名称と音の学習を小学校で行うことになったと解釈できるが、名称の学習に比べて、音の学習については明確に目標として設定されていない。さらに中学校以降での学習を考慮し、小学校では「音声と文字との関連付け」のみを行い、「発音と綴りを関連付けて指導」しないとされていることから、現場では文字と音との学習については明確な目標設定ができず、あまり積極的に取り組まれていないようである。

「書くこと」については、(ア)の目標では「語句や表現を書き写すことができる」と書かれているので、大文字、そして特に小文字を早く正確に書く力が必要となる。文字を認識し、書く力が十分に備わっていなければ、児童は書くことに苦痛を感じ、それらの活動には多くの時間がとられることになる。次に(イ)の目的は「例文を参考に書くことができる」である。ここでは「写す」ではなく「書く」とされている。つまり、学習者は自分がどのような語句を書いているのか十分に理解して書くことになり、音を文字化する（encode）力があることを意味する。「読むこと」に比べ、「書くこと」の目標には文字と音との関連を知る力についての言及がなされていない。しかし、筆者は文字と音との関係を意識しながら単語を写したり、または書いたりする力を身につけることは大変重要であり、書きながら音が生頭の中で再生できる力の育成がこれ以降のリタラシー獲得に大きな影響を及ぼすと考えている。

母語話者を対象とした初期リタラシー指導についてここで簡単にまとめておきたい。Chall（1967）は乳児から大人にいたるまでのリタラシー発達を0～6段階に分け、最初の3段階を「Learning to Read」、つまり読むこと自体を学ぶ期間とし、残りの3段階を「Reading to Learn」、読むことを通して学ぶ期間と捉えた。また、この境目は young adolescent の初めである小学校4年生としている。アメリカではより深い解釈（deep comprehension）ができる学習者を育てるため、Learning to Read の段階での指導が重要だとされ、研究が進められてきた。その中でもアメリカ議会の要請で招集された Reading Panel（全米調査委員会）ではそれまでに行われた実験研究などの結果をメタ分析し、小学校高学年までに獲得すべきリタラシーの基礎能力を phonemic awareness（音素認識能力）、oral language（音声言語）、phonics, vocabulary, fluency（流暢さ）とした。専門家や教師などで構成されたこの委員会では、教育現場でこれらの能力を意識的に育成することを推奨

している (National Reading Panel, 2000)。

英語は音と文字との関連が複雑なため、母語話者にとってもディコード (文字の音化) が難しい言語である。しかし、英語は基本的には表音言語なので、文字がどのような音も持つのかを理解し、単語がディコードできるスキルを獲得することは母語話者のみに限らず、外国語学習者にとっても大切であり、初期リタラシー指導の要である。その観点からすると学習指導要領の「書くこと」の指導目標に音と文字との関連を知り、単語を書くことができるスキルを身につけることが含まれていないのは、残念なことである。

2.2. 新学習指導要領に見られる評価の観点

2017年告示の学習指導要領では、小学校高学年の外国語教育においても「観点別学習状況の評価」として「知識・技能」、「思考・判断・表現」、「主体的に学習に取り組む態度」の3観点で評価を行うこととしている。また『「主体的に学習に取り組む態度」の評価に関しては… (中略) …知識及び技能を獲得したり、思考力、判断力、表現力等を身に付けたりするために、自らの学習状況を把握し、学習の進め方について試行錯誤するなど自らの学習を調整しながら、学ぼうとしているかどうかという意思的な側面を評価することが重要である』とされている。具体的には「① 知識及び技能を獲得したり、思考力、判断力、表現力等を身に付けたりすることに向けた粘り強い取組を行おうとする側面と、② ①の粘り強い取組を行う中で、自らの学習を調整しようとする側面」から評価することが求められている (国立教育政策研究所, 2022, pp. 9-10)。

また外国語教育の評価に関して文部科学省は、「国際共通語としての英語力向上のための5つの提言と具体的施策」(2011)において、中・高等学校は学習指導要領に基づき、生徒に求められる英語力を達成するための学習到達目標を「CAN-DO リスト」の形で具体的に設定することを提言した。2013年には、『各中・高等学校の外国語教育における「CAN-DO リスト」の形での学習到達目標設定のための手引き』(文部科学省, 2013)が発行され、CAN-DO リストの作成が推進されており、現在は小学校での外国語においても Can-Do リストの作成が奨励されている。

Can-Do Statements (これ以降 CDS) とは文字通り外国語のできる行動を記述した「能力記述文」である。小学生を対象とした CDS はまだ少ないが、CEFR-J の Pre-A1 レベルでは「読むこと」として「先生の読んでくれる絵本に出てくる英語の単語を見つけて選べる」と「ブロック体で書かれた大文字・小文字がわかる。」が発表されている (投野, 2013)。小学生を対象とした場合は、授業数や授業の目的などにもよるが、獲得するスキルの量や質も中・高校生とは異なり、多量のスキルを獲得するわけではない。アレン玉井 (2016) は下記のような CDS を提案した。

- ① 文字の形と名称 アルファベットの大(小)文字の名称を聞いてどの文字かわかる。
アルファベットの大(小)文字を見て、名称を言うことができる。
アルファベットの大(小)文字の名称を聞いてその文字が書ける。
- ② 文字の形と音 アルファベットの大(小)文字の音を聞いてどの文字かわかる。
アルファベットの大(小)文字を見て、音を言うことができる。
アルファベットの大(小)文字の音を聞いてその文字が書ける。

2.3 自己評価について

伝統的なテストによる能力測定に慣れている教師は、学習者の主観的な自己評価で学習者の本当の能力が測定できるのかという点について疑問を抱く。しかし、学習者自身が自分の学習をモニターし、必要な修正をするための情報を得、自己調整を有効に行うスキルを身につけるためにも自己評価は大切であり、学習の振り返りが推奨されている (Brown, 2007)。ここでは総括的な自己評価 (self-assessment of learning) と形成的な自己評価 (self-assessment for learning) について少し考えてみたい。

まずは、学習の成果を測定するための自己評価, self-assessment of learning であるが、学習成果を測定する総括的な評価 (summative evaluation) としての自己評価にあたる。大人を対象とした研究からは、自己評価と最終的な成績や教師の評価が概ね相関していることが報告されているが、それでも自己評価の項目の中には関連性の低いものもある。一方、研究自体が少ないのだが、子どもの学習者を対象とした自己評価に関する研究では、就学前の子どもたちの自己評価は非常にポジティブで、7～9歳から少し減少傾向になり、11～13歳に降下すると報告されている (まとめとして Butler, 2016)。ピアジェの発達理論を参考にしてこの傾向を考えると、自分の視点のみで外界を把握する、いわゆる ego centric な幼児は自分の能力を高く評価するが、自己中心性から脱却し、自分を客観的に見ることができるようになる7歳以降からは自己評価が厳しくなると解釈できる。また、子どもは「なりたい自分」と「本当の自分」を区別することができず、「なりたい自分」の姿で自分の能力を判断してしまうともいわれている。

一方、学習者が自分の学習状況を把握し、次の学習を組み立てていくのに役立つ形成的評価 (formative evaluation) としての自己評価 (self-assessment for learning) がある。近年、形成的評価として自己評価を使用することで、どのように学びが向上するのかについて関心を持つ研究者が増えている。

研究 1

3. 研究1の目的と方法

研究1では、児童が行う自己評価の信頼性について検証するため、文字学習の定着度を測定するテストの結果と自己評価を比較した。自己評価は self-assessment of learning, つまり総括的な評価として使われている。具体的にはリッカー尺度と Can-Do 尺度によって測定した自己評価との関連性について研究した。

3.1 参加者

本研究には2つのグループの小学6年生が参加している。1つのグループはリッカート尺度での自己評価を行ったグループであり、東京都内の公立小学校に通う6年生58名（男子23名、女子33名、不明2名）が参加した。彼らは小学校1年から6年まで週1回の英語の授業を受けており、2016年2月、最後の授業で文字に関するテストを受け、リッカート尺度による自己評価を行った。

もう一つのグループは Can-Do 尺度での自己評価を行ったグループであり、東京都内の公立小学校に通う6年生68名（男子32名、女子26名、不明10名）が参加した。1年から4年までは週1回の英語の授業を受け、5年のときは年間50回、6年では週2回の授業を受けており、2020年3月、最後の授業で文字に関するテストを受け、Can-Do 尺度による自己評価を行った。

どちらのグループの児童も筆者より特別なリタラシープログラムを受けており、最初に文字の字形とその名称を理解し書くことができるように指導された。次に音韻・音素認識能力を育てる活動を経験し、明示的、また体系的なフォニックス指導のもと、参加児童は全員文字と音との関係を意識的に学習した。どちらのグループでもリタラシー学習は毎回の授業で帯活動として7分ぐらいの時間をとって実施され、筆者と学級担任との Co-teaching で授業は進んだ。このプログラムは、文字の名前から文字の音の学習に移るときに行う音素体操と onset-rime に基づく、いわゆる new phonics が導入されているという特徴を持つ（アレン玉井, 2013）。

3.2 使用したテストと自己評価

アルファベットの文字と名称についての知識・スキルの測定のため、児童は4文字程度の名称を続けて聞き、それを書きとった。例えば「kei, el, eks, bi:」と2回聞いたのち、7秒以内に「klxb」と書くことになる。大文字と小文字ともに26文字を同様の形式で測定した。

次に音と文字との関係に関する知識・スキルを測るために、リッカート尺度のグループでは児童は音を聞いて相当する文字を書くというテストを受けた。例えば「/k/」と聞くと「c と k」を書いた（20 項目）。一方、Can-Do 尺度のグループでは、児童は言われた 3 つの単語に共通する最初の音を表す文字を書いた。例えば「bag, blue, bed」と 3 つの単語を聞き、児童は「b」を書いた（21 項目）。したがって 2 つのグループで使用された測定方法は異なる。

次に今回使用した自己評価方法であるが、1 つはリッカート尺度を使用したものである。内容としては文字とその名称、また文字とその音について (1) 理解できるかという receptive knowledge と、(2) 口頭で言えたり、書けたりするかという productive skill についての Can-Do Statements (CDS) について 4 件法で自分の力を評価する。下に例を示すようにそれぞれの知識・スキルについての CDS について、児童は「全然できない」から「とてもよくできる」という 4 段階で自分たちの力を評価した。

(例) 名称を聞いてどの文字かわかる。

全然できない あまりよくできない まあまあできる とてもよくできる
 |-----|-----|-----|

一方 Can-Do 尺度では、文字の名称と音に関する知識・スキルを問う CDS について、4 つのレベルに分けて書かれている記述文を読み、児童は該当するものを選んだ。下の例に示すように第一レベルの「名称を聞いてもどの文字かわからない」から第 4 レベルの「複数の文字の名称を早く言われてもどの文字が正確にわかる」まで、記述文を読み、児童は自分の知識・スキルに相当するレベルを選択した。記述文は授業の内容をもとに作成している（付録参照）。

(例) 文字の名称を聞いてどの文字かわかる。



これらのテストおよび自己評価の実施に先立ち、研究目的、ならびに個人情報匿名化され適切に管理されることを文書と口頭で校長および学級担任に説明し、参加の同意を得た。

4. 研究1のデータ分析と考察

ここでは測定したそれぞれのスキルと2つの尺度での自己評価を比較することにより、児童の自己評価の正確性について報告する。

4.1 名称を聞き大文字を書くスキルの測定と自己評価

リッカート尺度での自己評価グループと Can-Do 尺度での自己評価グループの大文字書き取りテストの得点をレベルごとに比較した。表1に記述統計を載せている。

表1 リッカート尺度と Can-Do 尺度のレベル分けによる大文字書きテストの結果

	人数	%	平均	中央値	標準偏差	四分位範囲
<u>リッカート尺度</u>						
レベル 2	1	1.9	*	*	*	*
レベル 3	17	31.5	22.7	24	3.41	6
レベル 4	36	66.7	24.6	25	1.52	2
<u>Can-Do 尺度</u>						
レベル 2	6	10.2	23.5	24.5	2.35	3
レベル 3	20	33.9	23.7	24	2.46	3
レベル 4	33	55.9	25.3	26	0.95	1

どちらの尺度でも、レベルが上がるにつれ平均点も上がっており、児童が自分の力を正しく評価していることを示唆している。次に、それぞれのレベルの得点到統計的な差があるかを調べた。レベル間に違いがみられるということは、児童はそれなりに自分の力を正しく評価していることを示す。各レベルの得点が正規分布をしていないので、Nonparametric の Kruskal Wallis 検定でレベル間を比較したところ、リッカート尺度でのレベル間には統計的に有意の差が見られなかった ($p = .103$) が、Can-Do 尺度ではレベル間に統計的な有意差が認められ ($p = .002$)、それはレベル2と4 ($p = .047$) とレベル3と4 ($p = .008$) の間の差であり、レベル2と3には差は見られなかった。名称から大文字を書くスキルに関しては Can-Do 尺度での自己評価にはレベル間に差が見られ、児童が正確に自分たちの力を評価していることがわかった。

4.2 名称を聞き小文字を書くスキルの測定と自己評価

小文字の書きとりテストについてもリッカート尺度での自己評価グループと Can-Do 尺度での自己評価グループをレベルごとに比較した。表2にその記述統計を載せている。

表2 リッカート尺度と Can-Do 尺度のレベル分けによる小文字書きテストの結果

	人数	%	平均	中央値	標準偏差	四分位範囲
<u>リッカート尺度</u>						
レベル 2	5	9.4	16.4	15	5.46	8
レベル 3	19	35.8	21.9	23	3.86	5
レベル 4	29	54.7	24.4	25	1.62	3
<u>Can-Do 尺度</u>						
レベル 2	10	16.9	21.8	22	2.53	3
レベル 3	18	30.5	23.7	24	2.24	3
レベル 4	31	52.5	24.2	25	3.82	2

どちらの尺度でも、レベルが上がるにつれ平均点と中央値が上がっており、児童が自分の力を正しく評価していることを示唆している。各レベルで統計的な差があるかを調べるため Nonparametric の Kruskal Wallis 検定で分析した。その結果、リッカート尺度でのレベル間で統計的に有意の差があった ($p = .007$) が、それはレベル 2 と 4 のみの差 ($p = .008$) であった。Can-Do 尺度でもレベル間に統計的な有意差が認められ ($p = .004$)、こちらもレベル 2 と 4 のみの差 ($p = .003$) であった。名称を聞いて小文字を書くスキルに関しては Can-Do 尺度およびリッカート尺度による自己評価にはレベル間に差があり、自己評価に信頼性があると考えられる。

4.3 文字の音とその小文字を書くスキルの測定と自己評価

最後に音素を聞いて小文字を書き取るテストについて、リッカート尺度グループと Can-Do 尺度グループをレベルごとに比較した。表3にその記述統計を載せている。

表3 リッカート尺度と Can-Do 尺度のレベル分けによる音一文字テストの結果

	人数	%	平均	中央値	標準偏差	四分位範囲
<u>リッカート尺度</u>						
レベル 1	2	3.8	10.5	10.5	7.78	
レベル 2	6	11.5	11.8	12	4.17	11
レベル 3	34	65.4	12.26	12	2.7	4
レベル 4	10	19.2	16.3	17	2.71	6
<u>Can-Do 尺度</u>						
レベル 1	2	3.4	13	13	2.83	
レベル 2	10	16.9	18.5	20	2.27	3
レベル 3	25	42.4	17.4	17	2.55	4
レベル 4	22	37.3	18.4	20	2.2	4

リッカート尺度では、レベルが上がるにつれ平均点も上がっており、児童が自分の力を正しく評価していることを示唆している。しかし、Can-Do 尺度ではレベル 3 の平均点がレベ

ル2よりも低く、平均点からはレベル2から4の間の違いが明確に見られない。

NonparametricのKruskal Wallis検定でレベル間を比較したところ、リッカート尺度でのレベル間で統計的に有意の差が見られた($p = .010$)が、それはレベル3と4のみ($p = .007$)であった。一方Can-Do尺度では、レベル間に統計的な差がある傾向が見られたものの、有意差は認められなかった($p = .069$)。前述したように音を聞いて文字を書きとるテストについては2つのグループが異なる様式のテストを受けたためにその信頼性を単純に比べることはできないが、レベル間に差が認められなかった一因にCan-Do尺度グループが受けたテストが簡単すぎて天井効果が出てしまったことが考えられる。またCan-Do尺度で使用していた記述文が児童の振り返りに役立つものであったのかを吟味する必要がある。

リッカート尺度とCan-Do尺度で自己評価を行ったグループのテスト結果を各レベルで分析した結果、大文字と小文字の学習については、Can-Do尺度を用いた自己評価が児童の実際の力をより正確に反映していることがわかった。しかし、音を聞いて文字を書くスキルに関しては、使用したテストが異なるため、単純に比較はできないもののリッカート尺度の信頼性がより高いものとなった。文字と音との関係に関するテストと自己評価の結果は、Can-Do尺度に用いた記述文の内容とテストが平易すぎたことによるのではないかと分析している。

しかし、Can-Do尺度に基づく4段階の自己評価では、教師は授業活動を反映させて記述文を作成することができ、また記述文を通してそれぞれのスキル向上に向けて児童や生徒にどのように学習目標を立てればよいのかを提示することができる。また学習者にとっては、記述文を読むことで自分の能力をより正確にまた具体的に判断することができ、次の目標を立てることができ、self-assessment for learning(形成的な自己評価)としても有効だと考えられる。

研究 2

5. 研究2の目的と方法

研究2では同じ小学校に通う5年生と6年生を対象に研究1で行ったCan-Do尺度による自己評価と実際の能力との関連について調査している。その目的の1つは、研究1の結果を再度検証することであり、もう一つの目的は年齢、また学習歴がどのように自己評価に影響を与えるのかを検証することである。

5.1 参加者

本研究には東京都内の小学校に在籍する5年生94名（男子40名，女子45名，不明9名）と6年生97名（男子41名，女子42名，不明14名）が参加している。参加児童は1年から4年まで週1回の授業，5年からは週2回の授業を受けていた。したがって6年生は5年生より約70回授業数が多くなる。彼らは2022年3月の最後の授業においてCan-Do尺度の自己評価と文字に関するテストを受けた。両学年とも研究1で説明したリタラシープログラムを受けており，授業形態は筆者を含め英語専科教員と学級担任とのCo-teachingであった。

5.2 使用したテストと自己評価

名称と文字についての知識・スキルについては，研究1で使用したテストを研究2でも使用した。しかし，音と文字との関係に関する知識を測るために，研究2では担当教員が発する音素を表す文字を3つの中から選ぶテストを実施した。例えば，教員が「/t/, /t/」と発音し，児童は紙面に書かれている「f, t, p」の中から「t」を選び丸で囲むというものである。音は2回繰り返され，問題の間には3秒おかれた。問題はすべて子音に関するものだが，6年生では20項目，5年生では18項目を出題した。

音と文字との関係に関するスキルを測るテストとして研究1と同様のものを使用した。児童は，3つの単語を聞き，最初の共通する音を表す文字を紙面に書く。例えば「bag, blue, bed」と聞き，「b」を書くというものであった。全て子音に関する問題だが，5年生では12項目，6年生では21項目出題した。

研究2では研究1でも使用したCan-Do尺度の自己評価を使用した。実際に使用したものは下記のようなものである。（付録参照）。

（例）音を聞いて文字を書く場合

(1)

文字は全然。
書けない

(2)

ゆっくり一文字
ずつなら書ける

(3)

複数の文字を速く
言われても、書ける

(4)

複数の文字を速く言われて
ても、正確に書ける。

研究2でもこれらのテストおよびアンケートの実施に先立ち，研究目的，ならびに個人情報匿名化され適切に管理されることを文書と口頭で校長および学級担任に説明し，参加の同意を得た。

6. 研究2のデータ分析と考察

最初にそれぞれの知識・スキルを測定したテストの得点と Can-Do 尺度の評価の関連性について、学年別に分析した。次に年齢および学習歴の違いにより、自己評価に変化があるのかを見ていった。

6.1 スキルの測定と Can-Do 尺度による自己評価の信頼性について

ここではそれぞれの学年で研究1同様、自己評価のレベル間でスキルの得点に統計的な有意差があるのかを調査した。まずは名称を聞いて大文字を書くスキルと自己評価を比較したところ、評価レベルが上がるにつれ平均点と中央値が両学年とも上がっているのが見られた。表4に記述統計を載せている。

表4 Can-Do 尺度のレベル分けによる大文字書きの結果

	人数	平均	中央値	標準偏差	四分位範囲	平均ランク
<u>5年生</u>						
レベル 2	19	20.95	22	4.37	6	28.58
レベル 3	48	23.33	25	3.1	3	43.91
レベル 4	13	23.38	25	3.33	4	45.35
<u>6年生</u>						
レベル 2	12	21.75	22	2.9	4	21.13
レベル 3	40	23.98	24	1.97	3	39.16
レベル 4	27	24.93	25	1.11	2	49.63

両学年ともにレベル1を選んだ児童はおらず、また全体の平均点から考えても大文字に関しての学習は十分に学習できているといえる。統計的には、データが正規分布していないので Nonparametric の Kruskal Wallis 検定で各レベルに差があるのかを分析した。

分析の結果、5年生ではレベル間に統計的な有意差がみられ、($p = .034$)、さらに詳しく調べるとレベル2とレベル3の間に差があることがわかった。6年生についてもレベル間に統計的な有意差がみられ ($p = .001$)、レベル2とレベル3、およびレベル2とレベル4の間に統計的な有意差がみられた。したがって大文字に関しては両学年ともに Can-Do 尺度で正確に自分の能力を評価しているようである。

次に名称を聞いて小文字を書くスキルとその自己評価について検証したところ、こちらも両学年ともに評価レベルが上がるにつれ平均点と中央値が上がっていた。表5に記述統計を載せている。5年生では9.5%、つまり約1割の児童がレベル1を選んでいるが、6年生では一人もレベル1を選んでいる。学年が進むにつれて小文字の学習が定着していることがうかがえるとともに、大文字に比べ小文字の習得には時間がかかることもわかる。こち

らも Nonparametric の Kruskal Wallis 検定で分析した。

表5 Can-Do 尺度のレベル分けによる小文字書きの結果

	人数	平均	中央値	標準偏差	四分位範囲	平均ランク
<u>5年生</u>						
レベル 1	8	14.5	16.5	5.83	11	11.75
レベル 2	22	19.77	20	3.99	6	29.8
レベル 3	41	22.95	24	3.38	3	50.28
レベル 4	13	23.92	25	2.93	3	58.38
<u>6年生</u>						
レベル 2	14	21.72	21	7.6	4	20.36
レベル 3	42	24.19	25	2.17	3	40.57
レベル 4	23	24.96	26	2.01	1	50.91

分析の結果、5年生ではレベル間に統計的な有意差がみられ、($p = .000$)、さらに詳しく見るとレベル1とレベル3、レベル1とレベル4、そしてレベル2と4の間に差があることがわかった。6年生についてもレベル間に統計的な有意差がみられ ($p = .000$)、レベル2とレベル3、およびレベル2とレベル4の間に統計的な有意な差がみられた。小文字に関しても両学年ともに Can-Do 尺度で正確に自分の能力を評価しているようである。

次に音を聞いてその文字を選ぶテストから測定した文字と音との関連を理解する力とその自己評価について分析した。5年生ではレベルが上がるにつれ平均点と中央値が上がる傾向にあり、音を聞いて文字がわかる力が身につけていると自己評価している傾向にある。一方6年生ではレベル間での得点に違いが見られず、テストに天井効果があり、平均点や中央値を見るとレベル3よりレベル2のほうが高い得点を示している。表6に記述統計を載せている。また5年生では7名、6年生では1名の児童がレベル1を選んでいった。例外が少ない子音と文字との関連であるが、その学習には小文字の学習同様にある程度の時間が必要であることがうかがえる。ここでもデータが正規分布していないので Nonparametric の Kruskal Wallis 検定で分析した。

表6 Can-Do 尺度のレベル分けによる音一文字選びの結果

	人数	平均	中央値	標準偏差	四分位範囲	平均ランク
<u>5年生</u>						
レベル 1	7	13.71	15	4.11	9	19.36
レベル 2	26	16.65	17	1.26	2	36.12
レベル 3	39	17.31	18	1.004	1	48.31
レベル 4	12	17.42	18.93	0.996	1	50.96
<u>6年生</u>						
レベル 1	1	*	*	*	*	*
レベル 2	15	19.53	20	1.64	2	38.1
レベル 3	39	19.51	20	1.39	3	36.73
レベル 4	24	19.96	20.5	1.43	2	44.88

分析の結果、5年生ではレベル間に統計的な有意差がみられ、($p = .004$)、さらに詳しく見るとレベル1とレベル3、レベル1とレベル4の間に差があることがわかった。6年生についてはレベル間に統計的な有意差がみられなかった ($p = .340$)。音を聞いて文字を選ぶ21点のテストで6年生は平均で19.61点と天井効果を示した。つまり、テストが簡単すぎたためにレベル間で差がつかず、自己評価の正確性を見ることができなかった。

最後に音を聞いて文字を書くスキルに関するテストとそのスキルに関する自己評価について検証した。表7に記述統計を載せている。5年生ではレベルが上がるごとに平均点と中央値が上がっているが、レベル4よりレベル3の平均値が高くなっている。6年生ではレベルがあがるごとに得点と中央値が上がっており、児童が自分の能力を正確に評価していたと思われる。ここでも Nonparametric の Kruskal Wallis 検定で分析した。

分析の結果、5年生ではレベル間に統計的な有意差がみられ ($p = .000$)、さらに詳しく見るとレベル1とレベル3、レベル1とレベル4、そしてレベル2とレベル3の間に有意差があった。前述したようにレベル4よりレベル3の得点が高いが、そこには有意な差はなかった。6年生についてもレベル間に統計的な有意差がみられ ($p = .002$)。レベル2とレベル3、およびレベル2とレベル4の間に統計的な有意な差がみられた。音を聞いて小文字を書くというスキルに関しては両学年ともに Can-Do 尺度で正確に自分の能力を評価しているようである。

表7 Can-Do 尺度のレベル分けによる音-文字書き結果

	人数	平均	中央値	標準偏差	四分位範囲	平均ランク
<u>5年生</u>						
レベル 1	11	6.36	7	4.34	9	18.36
レベル 2	31	10.06	10	1.26	2	36.11
レベル 3	32	11.13	12	1.31	2	54.69
レベル 4	10	10.9	11	1.29	2	49.85
<u>6年生</u>						
レベル 1	3	12.67	14	4.16	*	18.50
レベル 2	22	15.05	15	1.4	2	26.93
レベル 3	37	17	17	2.33	5	45.78
レベル 4	17	17.24	16	2.33	5	45.59

以上の分析結果をまとめると、名称を聞いて大文字また小文字を書くというスキルに関しては両学年とも分析の結果、Can-Do 尺度評価に信頼性があることが判明した。これは研究1と同様の結果であった。

一方、音と文字との関係については、音を聞いて文字を書くスキルに関しては、両学年とも Can-Do 尺度の自己評価が信頼できることを示していたが、音を聞いて文字を選ぶ知識に関しては、統一した結果がでなかった。5年生では Can-Do 尺度による自己評価の信頼性が示されたものの、6年生ではテストの平均点が高く天井効果を示したため、自己評価の信頼性を示すことができなかった。しかし研究1からも音と文字との Can-Do 尺度では十分にその信頼性を示すことができなかったことから、テストおよび Can-Do 尺度の記述文を変える必要があることが再確認された。

6.2 学年による自己評価の変化

5年生と6年生の Can-Do 尺度による自己評価を比較したところ、学年による差が見られた。そこで児童のレベル別の自己評価をそれぞれの知識・スキルごとに分けて学年別に回答を比較してみた。ここでは全体の有効回答数に対する各レベルの回答数をパーセンテージで示し、表8に載せている。下に結果がわかり易いように知識・スキルごとに学年ごとの変化がわかるようなグラフにしているが、白い棒は5年生、黒い棒は6年生の回答を表している。

表8 Can-Do 尺度による自己評価の学年差

	レベル1 (%)	レベル2 (%)	レベル3 (%)	レベル4 (%)
<u>大文字書き</u>				
5年生	3.6	22.9	57.8	15.7
6年生	0	15.2	50.6	34.2
<u>小文字書き</u>				
5年生	9.5	26.2	48.8	15.5
6年生	0	17.7	53.2	29.1
<u>音-文字読み</u>				
5年生	8.3	31	46.4	14.3
6年生	1.3	19	49.4	30.4
<u>音-文字書き</u>				
5年生	13.1	36.9	38.1	11.9
6年生	3.8	27.8	46.8	21.5

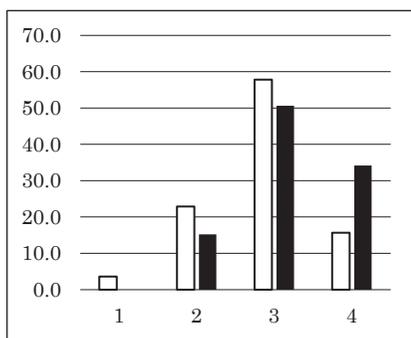


図1 大文字書きの自己評価変化

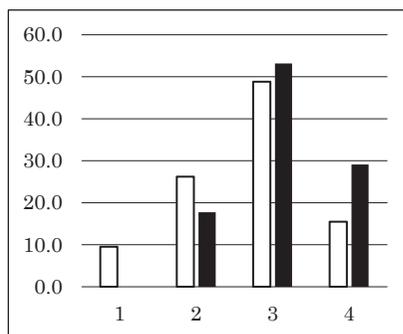


図2 小文字書きの自己評価変化

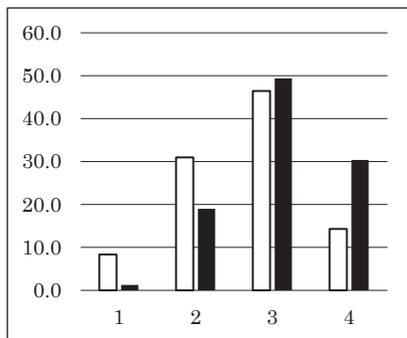


図3 音-文字読みの自己評価変化

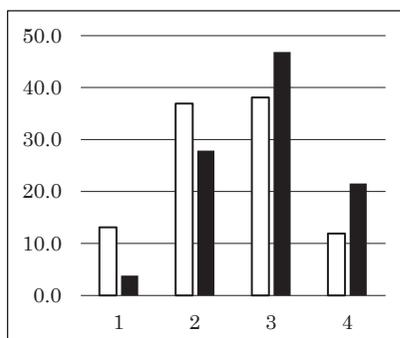


図4 音-文字書きの自己評価変化

6年生では名称を聞いて大文字、小文字を書くことに対して一番低いレベル「全然できない」を選んだ児童はいなかった。また、全体的に6年生の自己評価が高い傾向にあること

は、グラフからも一目瞭然である。音を聞いて文字を選んだり、書いたりする力に関しても、6年生の自己評価の方が高い。このように Can-Do 尺度による自己評価からは、児童自身が自分の力が着実に伸びていると評価している姿が見とれる。

7. まとめ

研究1では文字の理解や産出スキルについてリッカート尺度と Can-Do 尺度の自己評価とそれらの実際の力を関連して、自己評価の信頼性について検証した。その結果、名称を聞いて文字を書く力については、大文字、小文字ともにリッカート尺度および Can-Do 尺度による自己評価はそれなりに児童の実際の力を示唆し、とくに Can-Do 尺度ではその信頼性が高いことがわかった。一方、音を聞いて文字を書く力については能力を測定したテストが異なるため単純に2つの自己評価方法を比較することはできないものの、Can-Do 尺度による自己評価には十分な信頼性があるとは言えなかった。

研究2では研究1で使用した Can-Do 尺度の自己評価と文字に関するテストを使い、再現研究を行いその信頼性について再度検証した。また年齢または学習時間の違いで自己評価がどのように変化するのかも調べた。

その結果、名称を聞いてアルファベットの大文字、小文字を書くスキルに関しては、両学年とも児童が正確に自分の文字に関する知識・スキルを評価していることがわかった。つまり、Can-Do 尺度による自己評価の信頼性、また有用性が研究1同様に示されたことになる。児童は Can-Do 尺度の自己評価の記述文を読みながら、学習を振り返り自分の能力をより正確に把握することができたのではないかと考える。

一方、音と文字との関係については、研究1とは異なり、音を聞いてその文字を書くというスキルについては、両学年とも自己評価の信頼性が示されたが、音を聞いて文字を選ぶテストでは、6年生のデータからは Can-Do 尺度の信頼性を示すことができなかった。Can-Do 尺度の記述文をより適切なものに変更すること、そしてその能力を測定するテスト項目について再考する必要があると分析した。

また自己評価を学年別に比較した結果、6年生の児童がより高く自分の能力を評価しており、学習の深度が自己評価に適切に表されていることがわかった。年齢が上がるにつれ、自己評価は厳しくなる傾向にあることが報告されているが、この研究から Can-Do 評価の記述文に合わせて参加児童たちは自分の能力を正確に評価できていたことがわかった。1学年の差ではあるが、思春期を迎えるこの時期、必要以上に自己評価が低くなる年齢で、しっかり自分の能力を評価できていたことは重要である。

二つの研究から Can-Do 尺度の自己評価では児童が自分たちの文字学習スキルを正確に評

価していることがわかった。前述したように、自己評価、特に幼児や児童の自己評価は、彼らの本当の能力を表すものではないとする報告が多い。それは往々にして評価項目の曖昧性が評価の正確性を左右していることに起因している。評価項目が漠然としていれば、大人でも内容理解に困り、その解釈には幅がでてくるであろう。しかし、今回の研究から、児童は自分の能力をどのように評価すればいいのかが理解できれば、正しく自分の力を評価することができることが示唆された。授業活動を反映させた適切な Can-Do 尺度の記述文を作成できれば、児童が行う自己評価も信頼性の高いものになるであろう。

形成的な評価のための自己評価 (self-assessment for learning) のところで述べたように、自己評価は自律した学習者育成に効果的だと言われている。Butler (2016) は、実際に学習のための自己評価を効果的に行うためには、①指導しているような方法で評価を行うこと、②評価する活動を注意深く選ぶこと、③評価規準を柔軟に設け、学習目標に合わせて調整する。④自己評価をする際の規準を、学習者がしっかり理解できるように説明すること、などの重要性を指摘している。Can-Do 尺度による自己評価では授業活動を反映させて記述文を作成することができ、それにより児童に学習目標をより明確に提示でき、彼らがより適切に省察できるように導くことができる。今後も学習者の学びを支援する自己評価の在り方について研究を続けていくつもりである。

引用文献

- アレン玉井光江 (2013). 「公立小学校における Synthetic Phonics の実践—アルファベット知識と音韻認識能力の発達—」『ARCLE REVIEW』7号, 68-78.
- アレン玉井光江 (2016). 「小学校英語におけるリタラシー活動の Can-Do 評価」『小学校英語 Can-Do 評価尺度活用マニュアル別冊』68-82,
- 文部科学省 (2013) 『各中・高等学校の外国語教育における「CAN-DO リスト」の形での学習到達目標設定のための手引き』
- 文部科学省 (2018) 『小学校学習指導要領 (平成29年告示) 解説 外国語活動・外国語編』東京：開隆堂出版社
- 国立教育政策研究所 (2022) 『「指導と評価の一体化」のための学習評価に関する参考資料 小学校外国語・外国語活動』(7版) 東京：東洋館出版社
- 投野由紀夫 (編) (2013). 『英語到達度指標 CEFR-J ガイドブック』東京：大修館書店
- Chall, J. (1967). *Learning to Read: The Great Debate*. McGraw Hill.
- National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read*. Washington, DC: National Academy Press.

付録

名称を聞いてどの文字かわかる。(大文字・小文字)



文字を見て、その名称を言うことができる。(大文字・小文字)



名称を聞いて文字を書くことができる。(大文字・小文字)



音を聞いてそれに対応する文字がわかる。(子音)



文字を見てその音を発音できる。(子音)



音を聞いてそれに対応する文字が書ける。(子音)

