

# 学習者の自己評価の正確さに関する年齢要因の影響： 英語授業における小学4年生と2年生の自己評価の可能性

物井 尚子

千葉大学教育学部

## A Comparison of Perceived Competence in L2 for Fourth Graders and Second Graders in English

MONOI Naoko

Faculty of Education, Chiba University, Japan

本研究では、小学4年生102名、2年生97名の児童に、自己評価に関する質問紙調査を実施し、英語の授業における自己評価導入の可能性を検討した。調査では、参加者である小学生が自身の英語力に関する自己評価を一つの領域として判断しているか、また、年齢によって自己評価の的確さに違いがあるかに注目した。質問項目は授業カリキュラムと連動させることで、より具体的な質問内容を用意した。収集されたデータは、因子分析とラッシュ・モデルによる分析を行った。その結果、因子分析では4年生が1因子解、2年生が2因子解を得た。4年生は、2年生と比較して、英語力に関する自己評価にまとまりがあり、的確に測定している可能性が高い。また、ラッシュ・モデルを用いて、参加者の能力推定値を踏まえ、各質問の項目難度推定値を分析した結果、4年生が2年生よりも2倍のばらつきをもって各項目に回答していることが明らかになった。

In this study, a questionnaire survey was conducted on self-evaluation about one's English proficiency to 102 fourth graders and 97 second graders, and the possibility of introducing self-evaluation was examined in English lessons. The existent study paid attention to whether the participants evaluated their English proficiency as a single factor, and whether there was a difference in the accuracy or appropriacy of self-assessment depending on their age. The question items in the questionnaire were linked to the English curriculum to prepare for more specific questions. The collected data were analyzed by factor analysis and Rasch model analysis. The results showed that in the factor analysis, the fourth-grade students obtained a one-factor solution and the second-grade students obtained a two-factor solution. The fourth graders were more likely to evaluate their English proficiency more accurately than the second graders. The results also showed that in the Rasch model analysis, the fourth graders answered each item twice as variedly as the second graders.

キーワード：自己評価 (self-assessment)・認知されたコンピテンス (perceived competence)

小学校英語 (English for elementary school students)・外国語活動 (Foreign Language Activities)

### 1. はじめに

2017年に新学習指導要領 (文部科学省, 2017) が告示され、この2年間は、移行期間として各公立学校で2020年の全面実施に向けた準備が進んでいる。今回の学習指導要領により、小学校では教科としての外国語指導が5、6年生で開始されること、また、これまでの外国語活動が3、4年生に前倒しとなることから、小学校では4年間にわたって児童が英語に触れる機会を得ることになる。そのことに関連して、外国語の目標も従来とは一線を画し、外国語に関する関心・意欲・態度の涵養に加え、知識・技能の育成を求める指導目標となっている。以下に、新学習指導要領の外国語科の目標を引用する。

外国語によるコミュニケーションにおける見方・考え方を働かせ、外国語による聞くこと、読むこと、話すこと、書くことの言語活動を通して、コミュニケーションを図る基礎となる資質・能力を次のとおり育成することを目指す。(p. 137)

知識・技能が指導目標に加えられたこと、さらには、その中身について上述の内容が示されていることから、その「評価」をどのようにして他教科の水準に引き上げていくかは、これからの挑戦的な課題の一つとなる。文部科学省 (2016) は「小学校の高学年の外国語教育を教科として位置付けるに当たり、『評定』においては、中・高等学校の外国語科と同様に、その特性および発達段階を踏まえながら、数値による評価を適切に行うこと (p. 195)」と述べている。さらに、高学年の教科としての「観点別学習状況の評価」についても「知識・技能」、「思考・

連絡先著者：物井 尚子 (nmonoi@faculty.chiba-u.jp)

判断・表現」,「主体的に学習に取り組む態度」の3観点についてバランスよく行う必要があるとしている。その際、注意すべき点として、

必要な資質・能力を育成するための学びの過程を通じて、筆記テストのみならず、インタビュー（面接）、スピーチ、簡単な語句や文を書くこと等のパフォーマンス評価や活動の観察等、多様な評価方法から、その場面における児童の学習状況を的確に評価できる方法を選択して評価することが重要である。(p. 195)

多様な評価方法を導入することで、教室環境における児童の学習状況を的確な把握の重要性が指摘されている。本研究では「自己評価」を一評価方法として加えることが可能か、その場合、学習者の年齢を考慮する必要があるのかを研究課題とし、児童の英語力に関する自己評価の傾向を、学年を変えて調査するものである。

## 2. 先行研究

本節では、自己概念に関する研究からの示唆、学習者の年齢と自己評価に関する研究、の順に先行研究を整理する。

### 2.1. 自己概念に関する研究からの示唆

Alsaker and Kroger (2006) は、自己概念における加齢の影響を調査した先行研究から、小児期から青年期にかけて、自己に対する具体的かつ簡素な記述から、より差別化された複雑な評価へと変化することを指摘する。その一例として挙げられるのが、Harter (1979, 1982, 1985) の competence (有能さ) に関する研究である。Shavelson, Hubner and Stanton (1976) は、自己概念は様々な要素から構成され、それらは general/global self-concept (一般的な自己概念) を頂点にして階層的な構造を成していると主張するように、有能さを漠然と測定するのではなく、cognitive (「学習に対する<sup>1)</sup>」), social (「友人関係に対する<sup>1)</sup>」), physical (「スポーツに対する<sup>1)</sup>」), そして general self-worth (「全般的な自分の生き方に対する」(櫻井, 1983, p. 60)) という4つの領域に区分し、それぞれの認知された有能さを測定しようとしている点の特徴的である。Harter (1982) は小学3～6年生に対して The Perceived Competence Scale for Children を実施し、参加者はそれぞれの領域を別個のものとして差別化していることを確認した。

Nagai, Nomura, Nagata, Ohgi and Iwasa (2015) は愛知県内の小学校4校に通う日本人の小学1～6年生768名に対し、日本語で作成した the Children's Perceived Competence Scale を実施した。その結果、competenceの4つの下位尺度は .75～.87の範囲に収まる信頼性係数(内的一貫性)を示した。このことから、参加者である小学生は、有能さを漠然と評価するのではなく、4領域を差別化し、それぞれの領域での自身の有能さを測定していたことがわかる。これは、Harter

(1982, 1985) を支持する結果であった。興味深いのは、学年が上がるにつれて、性別の関係なく数値の減少が見られたことであり、Harter (1982)、さらには、Kokenes (1974)、Rosenberg (1979)、Furusho (2003) などの先行研究に沿った結果となった。就学児童については、自身の有能さを測定する際に、漠とした有能さではなく、それぞれの能力、先生からの評価などに基づいて、自己をより客観的に評価することが可能であることを裏付けている。

### 2.2. 学習者の年齢と自己評価に関する研究

橋本 (1983) は、自己評価は学習の目標を持たせるために最も効果的である、とその有用性を指摘している。また、Oscarson (1989) は、自己評価という作業が、学習者の学習に対する意識を高め、学習の目的を明確化でき、さらに心的側面の変容、自律性の促進につながるとしている。

ベネッセ総合研究所が調査した『第3回学習基本調査』(2001a, 2001b, 2001c) によれば、小学5年生2,402名、中学2年生2,503名、高校2年生3,808名に対して学習到達度テストと学習に関する意識・実態調査の2種類を実施した結果、中・高校生の客観的成績評価と主観的な評価の相関が  $.50 < r < .65$  という値を示した。一方、小学生においては、学力と自己評価の間に同程度の相関は確認されなかった。校種が上がるにつれて自己評価が正確になるという調査結果は跡部 (2012)、慶應義塾大学外国語センター (2011) でも確認されている。

Butler and Lee (1996) は、韓国の小学4年生70名、6年生81名を参加者とした、小学生という年齢層に特化した研究であるが、英語の授業中にタスクを実施し、その直後にそのタスクについて自己評価を行った場合、また、授業と関連しない一般的な英語の運用能力について自己評価を行った場合、いずれの状況下でも、6年生のほうがより正確性が高いことが確認された。筆者らは、9～10歳に満たない児童は、年上の学習者と比較すると自身の外国語運用能力の評価について正確さが劣ると結論付けている。また、この不正確さには、認知的、メタ認知的な発達要因に加え、社会的(児童の社会経験値など)、情意的要因(性格など)が複合的に関わってくると説明している(Butler, 2018a, 2018b)。これらの結果をもとに、次の研究課題を設定する。

## 3. 研究課題

本研究では、小学4年生と小学2年生に対して、これまでの英語の授業で学習した内容に関する自己評価を行わせた場合、

- 1) 両学年は、自身の英語力に関する評価を一つの領域として判断できているか。
  - 2) 上級生である4年生のほうが、自身の英語力についてできること、できないことの判断をよりの確に行うことができるか。
- の2点を研究課題とする。

#### 4. 調査方法

##### 4.1. 参加者

本調査における参加者は、千葉県内のA大学附属小学校に在籍する4年生児童105名と2年生児童105名のうち、欠席者を除く4年生102名、2年生97名の計199名である。

##### 4.2. 調査の流れ

本研究の調査の流れについて説明する。本調査実施に際し、事前に校長の承諾を得た。また、A小学校児童は入学時に教育的研究に関する承諾書を提出している。4年生は、2018年3月2日(金)の英語の授業を使用し、質問紙調査が行われた。当日は、年度末の授業ということで、1年間に学んだ学習内容に関するインタビュー・テストを授業者2名(ALTおよび筆者)それぞれが児童1名ずつを順番に別室に呼び、対面で行った。その間、教室ではクラス担任が主導し、①自己評価、②アルファベットの大文字・小文字の書き取り、③単語の認識に関する質問を順番に行った。①自己評価の調査に費やした時間は約10分程度であった。1組は9:50~10:30、2組は10:50~11:30、3組は11:35~12:15の40分ずつを使用して実施された。

2年生は、2019年7月9日(水)の英語の授業を使用し、質問紙調査が行われた。当日は、夏休み前の最後の授業ということで、挨拶、英語の歌を終えた後、クラス全員に①リスニング・テスト、②自己評価を実施した。②自己評価の調査に費やした時間は約10分であった。児童が、質問項目の理解に不安がある場合には、クラス担任、ALT、筆者が対応した。1組は9:00~9:40、2組は9:50~10:30、3組は11:35~12:15の40分ずつを使用して実施された。

##### 4.3. 自己評価に関する質問項目

本研究の自己評価に関する質問項目は、4年生には年間の英語の授業における学習内容をもとに8問が用意された(資料1)。質問には4件法を使用し、「はい」「どちらかと言えばはい」「どちらかと言えばいいえ」「いいえ」の選択肢のうち、自分の気持ちに最も近いものを一つ選ぶというものである。一方、2年生は夏休み前までの英語の授業における学習内容をもとに6問が用意され

た(資料2)。質問には4件法を使用し、「とてもそうおもう」「おもう」「あまりおもわない」「ぜんぜんおもわない」の選択肢のうち、自分の気持ちに最も近いものを一つ選ぶというものである。

##### 4.4. データ分析

分析については、2点の研究課題に基づき、因子分析を実施し(研究課題1に対応)、次にラッシュ・モデルを用いて分析を行った(研究課題2に対応)。因子分析に際し、IBM SPSS®(Statistics ver. 24.0)(SPSS, Inc., 2016)、ラッシュ・モデルにはWinsteps®(Version 3.91.0)(Linacre, 2015)を使用した。

#### 5. 結果と考察

本節では、調査結果を研究課題に沿って詳述する。

##### 5.1. 参加者は英語力に関する自己評価を一つの領域として判断できているか

本節では、自己評価に関する質問項目に対する児童の回答結果をまとめる。表1、2は、各学年の記述統計をまとめたものである。4年生の結果を見ると、①自分の名前を英語で言える、②今の気分を英語で言える、のように毎回の授業で行うやり取りに関する質問項目には4点満点に近い値が与えられた。また、④ハロウィンに出てくるお化けの名前はだいたい英語で言える、のように児童にとって印象に残るトピックについては高い数値が示された。

次に、2年生の結果を見ると、全体的に高い数値が確認された。また、②「じゅぎょうでならった どうぶつのなまえが言える」、⑤「好きなやさいをえいごで言える」という直前の学習で学んだ項目に関する回答がより肯定的な値として示された。一方で、⑥嫌いな野菜を英語で言える、という質問には、⑤と比較してやや低い値が示された。また、この質問には無回答の参加者が13名存在した。質問の意図としては、⑤でI like…、⑥でI don't like…の構文が使えるかを間接的に尋ねたかったのであるが、「嫌いな野菜は一つもないので、この質問には答えられないからどうしたらいいですか」と尋ねる児童が複数存在したことが無回答の理由である。

表1 4年生の回答結果(記述統計)

	①名前	②気分	③時間	④ハロウィンのお化け	⑤鍵の場所	⑥職業	⑦体の部位	⑧英語の動き
平均	3.70	3.57	3.12	3.35	2.77	2.97	3.05	3.07
標準偏差	.67	.68	.91	.78	1.03	.97	.95	.89

表2 2年生の回答結果(記述統計)

	①歌	②動物の名前を聞く	③動物の名前を言う	④気分	⑤好きな野菜	⑥嫌いな野菜
平均	3.24	3.43	3.14	2.95	3.56	3.21
標準偏差	.86	.77	.90	1.12	.79	1.05

次に、参加学年それぞれの質問調査紙の回答について因子分析（最尤法・プロマックス回転）を行った。4年生については、固有値1以上の1因子解を得た。累積寄与率は53.78%であった。因子分析の結果を表3に示す。それぞれ表の左端にある項目は第1因子の因子負荷量の絶対値の降順となっている。

表3 4年生の質問紙解答の因子分析結果（パターン行列）

	因子
⑦体の部位	.878
⑧英語の動き	.856
④ハロウィンのお化け	.793
⑤鍵の場所	.714
②気分	.700
⑥職業	.657
①名前	.623
⑨英語の動き	.593

注：.40以上の値は太字で示す。

全ての質問項目が .59 ~ .88 という高い因子負荷量をもって第1因子に寄与していることから、この因子を「英語運用能力の自己評価」と見ることができる ( $\alpha = .898$ )。参加者である4年生は、上述の質問項目で英語運用能力を一つの領域として判断できていると理解した。

次に、2年生についても質問紙の回答について因子分析（最尤法・プロマックス回転）を行った。2年生については、固有値1以上の2因子解を得た。累積寄与率は47.18%であった。因子分析の結果を表4に示す。それぞれ表の左端にある項目は第1因子、第2因子の因子負荷量の絶対値の降順となっている。

表4 2年生の質問紙解答の因子分析結果（パターン行列）

	第1因子	第2因子
④気分	.912	-.165
⑥嫌いな野菜	.501	.171
⑤好きな野菜	.414	.218
①歌	.383	.208
③動物の名前を言う	-.029	.779
②動物の名前を聞く	.068	.735

注：.40以上の値は太字で示す。

第1因子は、質問項目①、④～⑥の4項目から構成されている。①「えいごのうたを大きなこえでうたえます」、④「じぶんの今の気もちがえいごでいえます」、⑤「好きなやさいをえいごでいえます」、⑥「きれいなやさいをえいごでいえます」の質問項目が .41 ~ .91 という高い因子負荷量をもって第1因子に寄与していることから、この因子を「英語運用能力の自己評価」とみることができる。信頼性係数は  $\alpha = .681$  であった。第2因子は、質問項目②、③の2項目から構成されている。②「じゅぎょうでならった どうぶつのなまえをきいてわかります」、③「じゅぎょうでならった どうぶつのなまえがいます」

の質問項目が第2因子に寄与していることから、この因子を「動物に関する英語表現使用に関する自己評価」とみることができる ( $\alpha = .740$ )。

上記の結果から、参加者である4年生は、上述の質問項目で英語運用能力を一つの領域として捉えていると判断した。一方、2年生は英語に関する自己評価を一つの領域として判断していない可能性が確認された。

## 5.2. 4年生のほうが、自身の英語力についての判断をよりの確に行っているか

本節では、ラッシュ・モデルを使用し、自己評価に関する質問項目に対する児童の回答結果を分析した。その結果、4年生については、受験者能力推定値の平均は2.16で、標準偏差が2.29、受験者能力推定値の信頼性は .80 であった（表5）。また、項目難度推定値の平均は .00、標準偏差は1.08、項目難度推定値は .96 という結果が得られた。受験者能力推定値と項目難度推定値の対応を比較すると、項目難度推定値を受験者能力推定値が超えており、天井効果を示唆する結果となった。能力推定値が高い受験者に対する質問項目が少なかったために、受験者の能力値の推定には誤差が多く含まれている。McNamara (1996)、Bond and Fox (2007) の基準に基づき、受験者能力推定値のインフィット平方平均 (infit MNSQ) が .75 ~ 1.30 の範囲に収まっているかを確認したところ、ミスフィットに該当するのは102名のうち50名であった。これにより、能力推定値と質問項目のミスマッチが明らかになった。

表5 受験者能力推定値と項目難度推定値

	数	能力値 (上) / 項目 難度 (下) (標準偏差)	Infit MNSQ (標準偏差)	Outfit MNSQ (標準偏差)
受験者	102	2.16 2.29	1.02 .71	1.00 .80
項目	9	.00 1.08	1.02 .24	1.00 .24

また、図1は、受験者能力推定値と項目難度推定値を1つのグラフに収めたものである。これを見ると、4年生児童にとって「だれかのカギがつくえの上においてあったら、先生に英語でおしえられる」という前置詞の使用を求められる英文使用が、最も難易度が高いと判断されるが、この項目でさえも、参加者の能力推定値では平均的な値の近くに位置している。

次に、項目別に項目難度推定値、次元性の残差負荷 (residual loading)、インフィット平方平均 (infit MNSQ)、アウトフィット平方平均 (Outfit MNSQ) を見ていく（表6）。項目数が8問と2年生よりも2問多く、年度末に実施した調査であった。項目難度推定値が  $\pm 2.0$  の範囲に収まっている。これは2年生の項目難度推定値の約2倍のばらつきであり、彼らの自己評価の正確性が2年生を上回る可能性が高いことを示唆している。

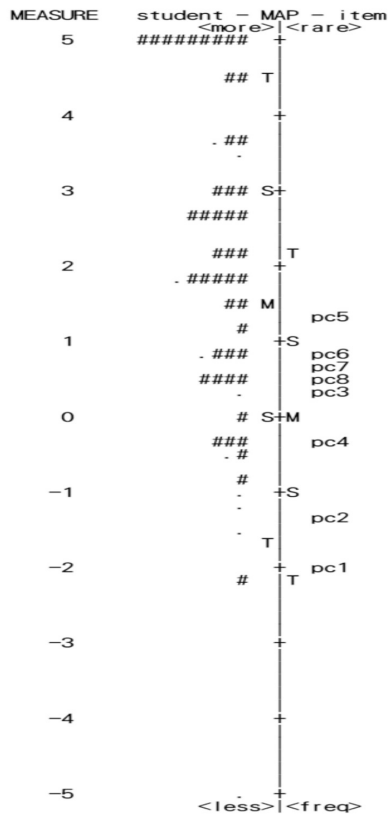


図1 受験者能力推定値(左)と項目難度推定値(右)の対応

表6 各項目の項目難度推定値, 次元性の残差負荷 (residual loading), インフィット平方平均 (Infit MNSQ) およびアウトフィット平方平均 (Outfit MNSQ)

	項目難度推定値	residual loading	Infit MNSQ	Outfit MNSQ
項目⑤	1.40	-.62	1.31	1.35
項目⑥	.83	.73	.69	.71
項目⑦	.59	.59	.75	.73
項目⑧	.55	-.41	1.36	1.38
項目③	.35	-.32	.84	.86
項目④	-.40	.02	1.00	1.02
項目②	-1.31	.18	1.00	.92
項目①	-2.01	.31	1.23	1.03

続いて、2年生の回答結果を同様の流れで分析した。受験者能力推定値の平均は1.64で、標準偏差が1.68、受験者能力推定値の信頼性は.62であった。また、項目難度推定値の平均は.00、標準偏差は.53、項目難度の推定値は.89という結果が得られた。項目難度推定値の標準偏差が、4年生と比較すると半減しており、ばらつきが少なかったことが分かる。さらに、受験者能力推定値と項目難度推定値の対応を比較すると、2年生と同様に、項目難度推定値を受験者能力推定値が超えており、天井効果を示唆する結果となった。能力推定値が高い受験者に対する質問項目が少なかったために、受験者の能力値の推定には誤差が多く含まれている。McNamara (1996),

Bond and Fox (2007) の基準に基づき、受験者能力推定値のインフィット平方平均 (infit MNSQ) が.75 ~ 1.30 の範囲に収まっているかを確認したところ、ミスフィットに該当するのは97名のうち40名であった。これにより、能力推定値と質問項目のミスマッチが明らかになった。

表7 受験者能力推定値と項目難度推定値

	数	能力値(上) / 項目 難度(下) (標準偏差)	Infit MNSQ (標準偏差)	Outfit MNSQ (標準偏差)
受験者	97	1.64 1.68	.98 .54	.97 .55
項目	6	.00 .53	1.00 .20	.97 .22

また、図2は、受験者能力推定値と項目難度推定値を1つのグラフに収めたものである。これを見ると、2年生児童は「自分の今の気持ちを英語でいえる」という問いに対し、最も難易度が高いと感じており、「好きなやさいを英語でいえる」という直近の学習内容に関する質問に対し、最も簡単であると感じていることが確認された。前者については、授業中に毎回行う“How are you?” “I am good/ fine/ happy.”等のやり取りであり、ほぼ全員の児童が毎回積極的にやり取りを行っている。回答の数値が低かった理由については、回答時、「今の気持ち」という日本語表現から“How are you?”を用いて行うやり取り表現と結びつかなかった可能性がある。

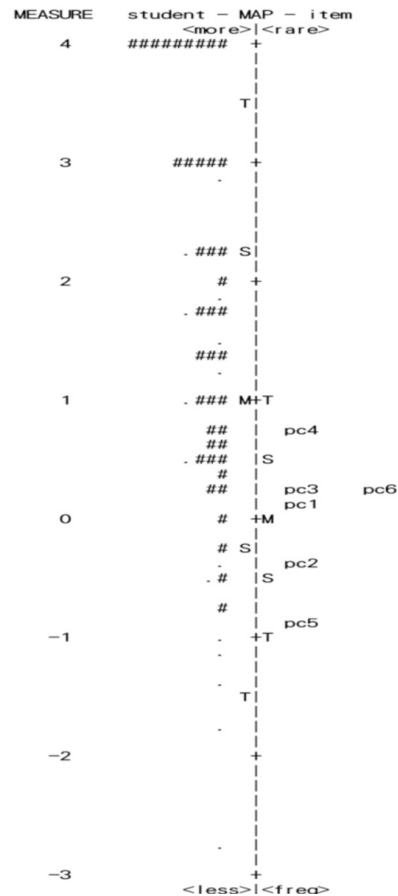


図2 受験者能力推定値(左)と項目難度推定値(右)の対応

次に、各項目別に項目難度推定値、次元性の残差負荷 (residual loading), インフィット平方平均, アウトフィット平方平均を見ていく (表8)。項目数が6問と4年生よりも数が少ないこと、4年生は年度末に実施した調査であったが、2年生は夏休み前の実施であったことを加味する必要がある。項目難度推定値が±1.0の範囲に収まっている。言い換えれば、2年生はどの項目についても、答えやすさに大きな違いはなく、類似した回答を寄せる傾向があることが明らかにされた。これは、彼らの自己評価の的確さが4年生と異なる可能性が高いことを示唆していると考えられる。

表8 各項目の項目難度推定値、次元性の残差負荷 (residual loading), インフィット平方平均 (Infit MNSQ) およびアウトフィット平方平均 (Outfit MNSQ)

	項目難度 推定値	residual loading	Infit MNSQ	Outfit MNSQ
項目④	.71	-.65	1.10	1.14
項目⑥	.28	-.19	1.25	1.06
項目③	.27	.81	.87	.91
項目①	.08	-.11	1.16	1.28
項目②	-.44	.70	.66	.62
項目⑤	-.90	-.34	.98	.80

## 6. 結論

今回、2つの研究課題を立て、4年生、2年生の児童に、自己評価に関する質問紙調査を行った。その結果、因子分析では、4年生において8項目の質問数で1因子解、2年生では6項目の質問数で2因子解を得た。このことから、4年生は自分の英語力を評価するような質問に対する回答に、よりまとまりがあり、自分自身の他教科や他分野の能力 (スポーツ・算数等) と差別化をはかりながら、英語力という技能を的確に測定する可能性が考えられる。Harter (1982, 1985) 等の先行研究を支持する結果となった。次に、ラッシュ・モデルを用いて、参加者の能力推定値を鑑み、各質問項目の項目難度推定値を分析したところ、4年生のほうが2年生よりも2倍のばらつきをもって回答していることが明らかになった。つまり、質問によって、自らのできること、できないことをよりの確に把握しており、それが回答のばらつきとして現れた可能性が考えられる。年齢に応じて、自己評価力に変化があることが確認された。

今回は、児童の学習カリキュラムに沿った自己評価項目を設定したこと、学年による実施時期がずれたことから、4年生が8項目、2年生が6項目と質問数に差があったことが調査にバイアスを与えてしまった可能性がある。また、学習内容が異なるので質問内容にずれがあるのはやむを得ないことではあるが、幾つかの質問を同じくすることで、それらの質問項目に対する学年別の自己評価の値を比較できたかもしれない。研究計画に更なる改善の余地がある。

この限界点を考慮しつつ、自己評価に対する学習年齢の影響について4年生という発達段階の鍵になる年齢と

より若い2年生を比較できたことは、大きな意味があると考えられる。また、教室環境での小学校英語教育に対する教育的示唆として、自己評価を何の目的で、どの学年に実施していくか、という実践的な問いに対する一つの答えを示したことは有意義であるといえる。今回の結果をもとに、今後の自己評価の在り方を考えていきたい。

## 7. 今後の課題

今後の研究では、複数の学年に対して調査を継続する必要がある。また、多様な地域の学校で調査を行うことが望ましいであろう。また、自己評価項目を精選した上で学年を越えて共通の項目を設定する、調査時期を一致させる、という2点を実現することで、より多様な年齢層を含む、包括的な自己評価の在り方を検討できることになろう。今後も英語という学習言語の学びにおけるより良い自己評価の取り入れ方を模索していきたい。

注1 日本語訳は全て櫻井 (1983) による。

## 謝辞

本研究は、平成28-30年度 科学研究費助成事業基盤研究 (C) (16K02954研究代表者 物井尚子) の支援を受けて行われた。ここに感謝の意を表する。

## 引用文献

- Alsker, F. D., & Kroger, J. (2006). Self-concept, self-esteem, and identity. In S. Jackson & J. Goossens (Eds.), *Handbook of adolescent development* (pp. 90-113). New York: Psychology Press.
- Bond, T., & Fox, C. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2<sup>nd</sup> ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Butler, Y. G., & Lee, J. (2006). On-task versus off-task self-assessment among elementary school students. *The Modern Language Journal*, 90(4), 506-518.
- Butler, Y. G. (2018a). The role of context in young learners' processes for responding to self-assessment items. *The Modern Language Journal*, 102(1), 242-261.
- Butler, Y. G. (2018b). Young learners' processes and rationales for responding to self-assessment items: Cases for generic can-do and five-point Likert-type formats (pp. 21-39) In Davis, J. et al (Eds.) *Useful assessment and evaluation in language education*. Georgetown, NY: Georgetown University Press.
- Furusho, J. (2003). A study of the Kid-KINDL questionnaire for measuring quality of life in elementary school children in Japan. *The Journal of the Japan Pediatric Society*, 107(11), 1521-1526.
- Harter, S. (1979). *Perceived competence scale for children* (manual). Denver, CO: University of Denver.
- Harter, S. (1982). The perceived competence scale for children. *Children Development*, 53, 87-97.

- Harter, S. (1985). The Self-perception profile for children: Revision of the perceived competence scale for children (manual). Denver, CO: University of Denver.
- Kokenes, B. (1974). Grade level differences in factors of self-esteem. *Developmental Psychology*, 10 (6), 954-958.
- McNamara, T. F. (1996). *Measuring second language performance*. London: Longman.
- Nagai, Y., Nomura, K., Nagata, M., Ohgi, S., & Iwasa, S. (2015). Children's perceived competence scale: Reference values in Japan. *Journal of Child Health Care*, 19 (4), 532-541.
- Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and implications. *RELC Journal*, 19 (1), 75-93.
- Rosenberg, M. (1979). *Conceiving the self*. New York: Basic Books.
- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46 (3), 407-441.
- 跡部智 (2012) 「小中高大連携目指す共通枠を活用した AOP の報告と課題」『第21回 JACET 春季英語教育セミナー 小中高大連携の英語教育を目指す「英語共通枠」活用の試み (発展編)』口頭発表資料. 2012年3月11日早稲田大学.
- 慶應義塾大学外国語教育研究センター (2011) 『行動中心複言語学習プロジェクト Action Plurilingual Language Learning Project (AOP) 平成18~22年度私立大学学術研究高度化推進事業 (学術フロンティア推進事業) 研究成果報告書』横浜市: 慶應義塾大学出版会株式会社.
- 櫻井茂男 (1983) 「認知されたコンピテンス測定尺度 (日本語版) の作成」『教育心理学研究』31, 245-249.
- 橋本重治 (1983) 『新・教育評価法総説』下巻 東京: 金子書房.
- ベネッセ総合教育研究所 (2001a) 『第3回学習基本調査報告書・高校生版』東京: —.
- ベネッセ総合教育研究所 (2001b) 『第3回学習基本調査報告書・小学生版』東京: —.
- ベネッセ総合教育研究所 (2001c) 『第3回学習基本調査報告書・中学生版』東京: —.
- 文部科学省 (2017) 『小学校学習指導要領 外国語活動・外国語編』東京: —.

資料1 4年生用英語力自己評価項目

- ① 自分の名前を英語でいえる。  
ア. はい イ. どちらかと言えばはい ウ. どちらかと言えばいいえ エ. いいえ
- ② いまの自分の気分を英語で言える。  
ア. はい イ. どちらかと言えばはい ウ. どちらかと言えばいいえ エ. いいえ
- ③ 英語で時間を言われたら、わかる。  
ア. はい イ. どちらかと言えばはい ウ. どちらかと言えばいいえ エ. いいえ
- ④ ハロウィンに出てくるお化けの名前はだいたい英語で言える。  
ア. はい イ. どちらかと言えばはい ウ. どちらかと言えばいいえ エ. いいえ
- ⑤ だれかのカギがつくえの上においてあったら、先生に英語でおしえられる。  
ア. はい イ. どちらかと言えばはい ウ. どちらかと言えばいいえ エ. いいえ
- ⑥ いろんなしよくぎょうを英語で言える。  
ア. はい イ. どちらかと言えばはい ウ. どちらかと言えばいいえ エ. いいえ
- ⑦ からだのいろんな部分を英語でいえる。  
ア. はい イ. どちらかと言えばはい ウ. どちらかと言えばいいえ エ. いいえ
- ⑧ 英語でいろんな動きをするように言われたら、その動きができる(サイモン・セズのことです)。  
ア. はい イ. どちらかと言えばはい ウ. どちらかと言えばいいえ エ. いいえ

資料2 2年生用英語力自己評価項目

- ① えいごのうたを大きなこえでうたえます。  
(とてもそうおもう・おもう・あまりおもわない・ぜんぜんおもわない)
- ② じゅぎょうでならった どうぶつのなまえをきいてわかります。  
(とてもそうおもう・おもう・あまりおもわない・ぜんぜんおもわない)
- ③ じゅぎょうでならった どうぶつのなまえがいえます。  
(とてもそうおもう・おもう・あまりおもわない・ぜんぜんおもわない)
- ④ じぶんの今の気もちがえいごでいえます。  
(とてもそうおもう・おもう・あまりおもわない・ぜんぜんおもわない)
- ⑤ 好きなやさいをえいごでいえます。  
(とてもそうおもう・おもう・あまりおもわない・ぜんぜんおもわない)
- ⑥ きらいなやさいをえいごでいえます。  
(とてもそうおもう・おもう・あまりおもわない・ぜんぜんおもわない)