

カリキュラムに基づく尺度の日本語版開発に向けた 算数の問題と採点手続きの検討

干 川 隆

Investigation of Math Problems and Scoring Procedure in Curriculum-based Measurement: Toward the Introduction of Standards in Japan

Takashi HOSHIKAWA

(Received October 1, 2014)

This study investigated math problems and scoring procedure in Curriculum-Based Measurement (CBM), with an aim toward introducing standards in Japan. In the United States, CBM is useful for monitoring the progression of students' learning within the Response to Intervention (RTI) movement. In the computation of CBM score, the scoring procedure was to award points for each digit in the correct place and for identification of correct units in the answer. The points were summed to give the total score. Two studies as part of this investigation were conducted. In the first study, 44 university students were asked to answer difficult and easy math problems and the basis of assessment was the total score. In the results, participants performed better on the easy problems, both in terms of the total score and the subscore for correct digits. In the second study, 38 university students were asked to solve two different types of problems that had the same difficulty, and they were assessed based on the total score. Results indicate that there were no significant differences between the two types of problems. These results are discussed with respect to the introduction of standards in Japan and the necessity of computer applications and consultation services for CBM.

Key words : Curriculum-based measurement, scoring procedure, mathematics, standardization, progressive monitoring

I. 問題と目的

1. RTI とそれを支える CBM

1) 学習障害の認定を巡る動向

アメリカ合衆国（以下米国）では、わが国と同様に特殊教育の見直しが行われ、2002年に特殊教育におけるエクセレンスに関する大統領諮問委員会（President's Commission on Excellence in Special Education）の報告書「新しい時代：子どもとその家族のための特殊教育の復興」が出された（U.S. Department of Education, 2002）。その諮問した背景の中には、①特異な学習障害者のうち、80%は単にどのように読むかを学習していないだけであること、②マイノリティにある子どもが特殊教育のいくつかのカテゴリーに過剰に表されていること、などが指摘されている。この委員会の出した勧告は、勧告1：過程よりも結果に注目しなさい、勧告2：失敗のモデルでは

なく予防的なモデルを取り入れなさい、勧告3：障害のある子どもを通常教育の子どもとしてまず考えなさい、であった。勧告2では、現在のモデルが子どもが失敗するまで待たなければならないモデルであるのに対して、科学的な基礎をもつ指導と教授方法を用いて、早期の認定と迅速な支援に向けてシステムを動かすことが提案されている。勧告3では、これまでの通常教育と特殊教育という二重の会計システムの問題点が指摘された。

この大統領委員会の勧告を受けて、2004年に障害のある人の教育法（Individuals with Disabilities Education Acts）が改訂された。その中には、特異な学習障害の児童生徒を認定する方法において重大な変化を導く条項が含まれていた（IDEIA, 2004）。それは、州によって採用された基準が、①子どもが学習障害をもつかどうかを決定するために、知的能力と学力の間の重度のディスレパンスの使用を必要としない、②科学的で、研究に基づく介入への子どもの応答によ

るプロセスを使用できなければならない、と規定されたことである。この規定は、これまで学習障害の認定方法として問題が多い（例えば、千川、1999）にもかかわらず、運用上の利点からほとんどの州で採用されていたディスレパシーモデルを排除し、「介入への応答（Response to Intervention：以下 RTI）」へと大きく方向転換をしたことを示すものである（Grigorenko, 2009）。

2) RTI

RTI は、単に学習障害の認定だけでなく、卓越した教育改革の一つとして位置づけられている（Gilbert, Compton, Fuchs, Fuchs, Bouton, Barquero, & Cho, 2013）。RTI は学習障害のある児童生徒に限らず、学力にリスクのある児童生徒の成績を改善することをねらった通常教育の予防システムにまで発展してきた（Kratowill, Volpiansky, Clements, & Ball, 2007）。加えて、予防的なプログラムがより大きな効果性をもち、指導が開始される前に問題が発生するのを防ぐ予防プログラムとしてコスト面で効果的であることが指摘されている（例えば、Torgensen, 2002）。RTI の特徴は、多層によるシステムである。主に3層での支援体制が確立され、第1層では、通常の学級においてユニバーサルなスクリーニングを行い、個々の学業の成長がモニターされる。そこでクラスメイトよりも著しく低いリスクのある児童生徒は、第2層へと移される。第2層では予防的な小グループの指導を受け、その進捗状況はモニターされる。第2層に反応せず効果的な指導から利益を得られなかった人は第3層へと移る。第3層ではこれまでの特殊教育と同様に集中した個別化された指導を実施する（Gilbert et al., 2013）。

RTI では、これまでの一度限りの（high-stakes）のアセスメントではなく、介入がどのように学力を押し上げているかを継続的に評価する必要がある。そのため、RTI では進捗状況をモニターするための手立てが用いられている。海津（2006）は、進捗状況のモニタリングによって、①子どもの学力の伸びに注目、②環境（教法や教材等）に関する問題と子ども自身のつまずきとの識別が可能、③指導を計画する上での手がかりの獲得、④指導が効果的であったか否かについての教員自身の評価の推進、のメリットを上げている。Fuchs, Mock, Morgan, and Young（2003）は、RTI の基本的な要素の一つは、スクリーニングと進捗状況モニタリングのための技術的に十分なシステムであると述べている。RTI ではその多くで、進捗状況をモニターするためにカリキュラムに基づく尺度（Curriculum Based Measurement：CBM, Deno, 1985）が用いられてきた。わが国では、進捗状況のモニタリングとして、わずかに海津・田沼・平木・伊藤・Vaughn（2008）に

よる多層指導モデル（MIM）があるに過ぎない。

2. CBM とは何か？

児童生徒の学力の進捗状況を把握するためには、これまで商業的な基準に準拠した学力テストが用いられてきたが、教えられたカリキュラムとテストされる内容とが必ずしも一致しないとの指摘がある（Jenkins & Pany, 1978; Shapiro & Derr, 1987）。また、教師が児童の進捗を評価するために児童の学力の非公式化な観察に頼っており、実際の児童の学力と児童の学力の教師の判断との間に乖離があり、学力テストの結果が指導に生かされていないことが指摘されている（Salmon-Cox, 1981）。教師は主に非公式な観察に頼ることが多いが、その際の信頼性や妥当性が問題となる。

1978年、連邦教育局は全米の5大学に学習障害の研究を委託し、ミネソタ大学学習障害研究施設（Institute for Research on Learning Disabilities）には、1983年まで学習障害の評価（assessmentとevaluation）を委託した（千川、2000）。その中で、ミネソタ大学のStaley L. Deno博士を中心とする研究グループはCBMを開発した。CBMは、読みでは、基礎リーダーからランダムに選択された教科書の文章を1分間声を出して読み、間違えずに生徒が読んだ単語の数を数える方法である。また、算数では、学年と一致した教科書から引き出された計算問題の2分間での正答だけでなく、計算の途中で正しい位置に書かれた答えの数を数える方法である。

CBMはよく、体温や血圧など日常にある体温計や血圧計などで測定できる「バイタルサイン」に例えられる（Deno, 1985）。つまり、教師は単純で妥当性があり効果的な手続きとしてCBMを用い、もしも平均からのレベルや個人内での進捗が基準に達しないとすれば、教え方や本人の学習に何らかの課題があるとしてさらに精密検査をしなければならない。バイタルサインとしてのCBMを定期的実施することによって、学習のつまずきについて追加の検査が必要なかどうかを教師が判断する物差しをもつことができるという利点がある（Deno, 1985; Deno & Fuchs, 1987）。CBMは1分間で実施でき、定期的に児童生徒の学習の進捗状況をモニターすることができる。また、これまでの研究からCBMは単に読みの能力だけでなく、文章理解などの全体的な国語能力を示すことが実証的に示されてきた（Marston, 1989; Deno, 1985）。

3. CBM のもつメリット

CBMを標準化することによって、次のようなメリットが考えられる。それはCBMによって、①特別支援教育と通常教育の場による指導の効果を比較検討

できること、②ユニバーサルデザインの授業づくりなどの新しい指導方法が提案されたときにその指導の効果を客観的に把握できること、③学習障害の児童生徒の通級による指導の対象となるのか指導を終了するのかの評価基準ができること、である。

①について、米国では研究者の多くはインクルージョンへの流れの中で、特殊教育と通常教育との指導の効果の違いを CBM を用いて比較してきた。その研究の多くは、特殊教育による指導の効果を示すものであった（例えば、Fuchs, Fuchs, & Fernstrom, 1993; Fuchs, Roberts, Fuchs, & Bowers, 1996）。Marston (1988) は、時系列の変化から特殊教育の効果について比較した。その結果、特殊教育としてリソースルームで教えられたときに、生徒の成績は2倍の速度で進捗したことを示し、11人の軽度障害のある子どものうち10人に効果が示されたことを明らかにした。したがって、CBM を標準化することができれば、指導の場の違いによる効果を場を越えて評価することができる。

②について、CBM は学習障害に限らず通常の学級にいる児童生徒の進捗状況をモニターすることができる。進捗状況をモニターすることができれば、指導の途中で不適切な介入方法を修正することや最終目標を修正することができる。今日、ユニバーサルデザインの授業づくり（廣瀬・桂・坪田, 2009）が実施され、通常の学級の中で発達障害のある児童も含めて、授業方法が工夫されるようになった。しかし、ユニバーサルデザインの授業づくりを導入することで生じる学習の評価は、必ずしも十分ではない。新しい介入方法の効果を評価する際に CBM のような評価方法がわが国でも標準化できると、その方法の効果をより客観的に把握することができるであろう。

③について、CBM のような全体成果尺度が作成できれば、指導の効果により特別支援教育の措置の判断に活用することができる。現在、わが国においては通級による指導の対象を決定する基準が明確でない。多くの判断は、就学指導委員会に任せられており、市町村や通級指導教室でその判断がまちまちである。もしも日本版の CBM という基準を作成することができれば、学習のつまずきのレベルと成長の二重の乖離 (dual discrepancy) によって児童生徒を評価することができる。したがって CBM の二重の乖離によって、対象児が通級による指導の対象なのか、また指導により通級による指導をやめるのかを判断することができる。

4. わが国における CBM の必要性

CBM はこのようなメリットがあるにもかかわらず、

わが国ではほとんど紹介されてこなかった。ではなぜこれまでわが国では、CBM が注目されてこなかったのでしょうか。米国では日本の学習指導要領のような National Curriculum Guideline がなく、児童生徒が当該学年の学習内容を習得したかどうかを判断するために、標準学力検査が発達してきた。一方、わが国では学習指導要領に基づき教育課程が編成され、それに対応する形で教科書が作成され、その学年の内容を習得したかどうかは、教科書の単元ごとの単元末テストや期末テストを評価することで判断されてきた。商業的なテストやドリルは主に単元末テストや期末テストが中心であり、その単元で習得すべきスキルの評価が行われてきた。しかし、米国の CBM の取り組みのように通常教育と特別支援教育（例えば通級による指導）の教育成果を比較するためには、単元ごとの質的な評価ではなく CBM のような量的な全体成果尺度が必要である。

わが国では、わずかに美坂 (2006) が CBM の標準化を試みてきただけである。美坂は、算数では計算を国語では視写を用いて、月に1度の児童生徒の進捗状況をモニターした。その結果、算数の計算と視写ともに発達的な変化を示すことができ、進捗状況をモニターするための方法として有用であることを示唆した。しかし、美坂の算数の採点方法は、児童が採点することもあり、問題の正答数のみで評価を行ってきた。月に1度の割合での評価であれば、問題の正答数でも十分に進捗状況を把握できるが、RTI のように教師の教え方が妥当なのかどうかを把握するためには、週による進捗状況を把握するために、さらに目の細かな成長に対して感受性の高い採点方法の開発が必要である。

一方、米国の算数の CBM は、単に正答数だけでなく結果としての答えが間違っていたとしてもそのプロセスの中で位置と数字が合っていた場合には、それにポイントが与えられる。本研究では、美坂の用いた問題をさらに検討し、米国で用いられている採点方法を採用するために基礎的な資料を得ることを目的とした。

以下に米国で用いられている CBM の採点例を示す。正答は、 $11 \times 91 = 1001$ であり、以下の例では誤答である。しかし、計算の途中で、数字と場所が正しいところ（下線部）ではそれをポイントとして換算することができる。したがって、この問題を正答したときには、全部で8ポイントを得ることができる。以下の筆算は誤答であるにもかかわらず、5ポイントを得ることができる (Tindal & Marston, 1990)。

$$\begin{array}{r}
 11 \\
 \times 91 \\
 \hline
 11 \longleftarrow 2 \text{ ポイント} \\
 19 \longleftarrow 1 \text{ ポイント} \\
 \hline
 1911 \longleftarrow 2 \text{ ポイント}
 \end{array}$$

5. 本研究の目的

前述のように、わが国でも通常の学級の中で多様なニーズのある児童生徒が増える中で、今後、米国のような RTI モデルをわが国の学校の教育システムに取り入れていくことが必要である。しかし、その際に具体的な手立てとして、CBM のように教師が容易に用いることができ児童生徒の進捗状況をモニタリングできる方法の確立が必要である。そこで本研究では、わが国で CBM を標準化するために、算数の問題の検討と米国で用いられているポイントによる採点方法の妥当性について検討することを目的とした。

II. 研究 1

1. 目的

研究 1 では、計算の CBM をわが国で標準化するために、CBM の得点を規定している要因を把握し、問題の違いによる差異を無くすために、美坂 (2006) が作成した小学校 6 年生の 7 題の計算問題の中から、米国の CBM の採点によって一番難しい問題と一番やさしい問題を選び、その間に差があるかどうかについて検討することを目的とした。

2. 方法

1) 被検者

被検者は、研究に協力の得られた大学生と専攻科学生 44 人 (20 歳から 49 歳まで、平均年齢 27.4 歳：女性 33 人、男性 11 人) であった。

2) 算数の問題

本研究では、美坂 (2006) が作成した小学校 6 年生の問題を用いた。美坂は 6 年生の問題を作成するにあたって、問題の 60% が当該学年の 6 年生の問題となり、残りを下学年の問題 (各学年ごとの問題を 3 年生から 5 年生まで、10%、10%、20%、計 40%) に成るように作成した。当該学年の問題を 60% 入れているのは、当該学年の学習が進行するにつれて正答できる問題が増え、学年終了時には、成績が右肩上がりのグラフに示せることを考えたからであった。美坂は 7 つの問題をこの割合に基づいて作成したが、問題間に難易度の違いがあると推測された。そこで、正誤だけでなく米国式に途中の場所と数字があっているものもポイントとして換算する採点方法を用いて 7 つの問題を評価した。その結果、米国のポイント換算方法で一番得点の高い問題 (研究 1 では No.1) と一番得点の低い問題 (研究 1 では No.2) を選択して検討することにした。なお、No.1 の問題数は 72 問、ポイント総数は 563 であり、No.2 の問題数は 72 問、ポイント総数は 515 であった。

3) 手続き

No.1 と No.2 の問題は、A4 版の用紙 1 ページに 18 問ずつ計 72 問を A3 版用紙に両面印刷で 1 枚にまとまるように印刷した。授業の終了した時間に問題を被験者に配布し、1 枚目の問題を 3 分間でできるだけ速く間違えずに解答するように教示した。難しい問題については飛ばしても構わないと教示した。実施後、5 分間の休憩の後に、2 枚目の問題でも 3 分間でできるだけ速く間違えずに解答するように教示した。

No.1 と No.2 の問題の順番は、解答への順番の影響を減らすために、半数の学生には No.1 の問題から、残りの半数は No.2 の問題から実施し、2 回目の問題でそれぞれの問題を入れ替えた。

採点の仕方は、美坂の用いた正答数と米国の手続きに基づくポイントの総数による二つの算出方法を用いた。美坂の方法では、最終的な正答数によって評価した。米国の CBM の方法では、最後まで正答していなくても途中の計算の場所と数字が合っていれば、ポイントが与えられた。途中の計算を省いて正答した場合でも、途中の計算をしているものと見なして、ポイントを算出した。被験者には、二つのテスト終了後に採点例を示した用紙を配布し、採点の仕方を説明して各自に採点させた。採点后、解答用紙を回収し、筆者が指示通りに採点しているかをチェックし、誤って採点したところは、指示通りになるように筆者が採点をしなおした。

4) 評価方法

CBM の評価には次の手続きを用いた。

- ①正答数の数
- ②途中の計算のポイントも数えたポイント総数
- ③ポイント総数による被験者内の二つの問題の相関関係

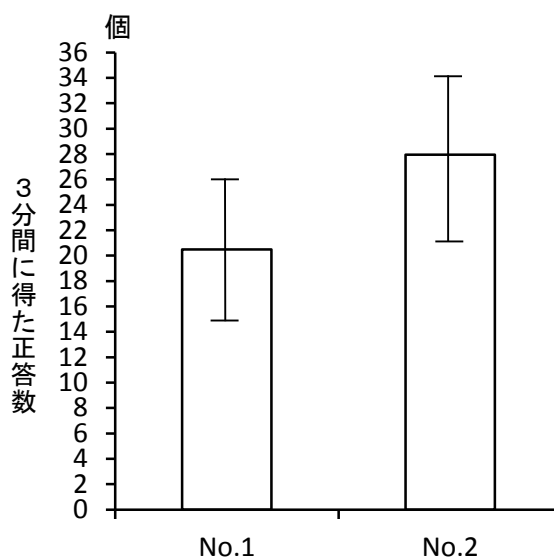


図 1. 正答数による No.1 と No.2 の問題の違い

④ポイント総数による提示順序による違い

3. 結果

1) 問題の正答数による比較

美坂の用いた正答数による比較の結果として、問題ごとの平均と±1標準偏差を、図1に示した。No.1とNo.2の問題の違いについて、t検定を行ったところ、No.1の問題に比べてNo.2の問題の正答数が有意に多かった ($t(43) = 9.61, p < 0.001$)。

2) ポイント総数による比較

米国式のポイント総数の評価によるNo.1とNo.2の結果は、図2に示した。図2に示すように、No.1の問題に比べてNo.2の問題のポイント総数が有意に高かった ($t(43) = 5.73, p < 0.001$)。

3) 分布の正規性の検討

分布の正規性について、No.1とNo.2それぞれの問題の正答数とポイント総数について尖度と歪度を求めたところ、いずれも有意ではなく、正規性は棄却できなかった。ポイント総数による度数分布については、図3に示す通りである。

4) No.1とNo.2のポイント総数の個人内の相関

被験者は、5分の休憩を挟んでNo.1の問題とNo.2の問題を実施した。したがって、両者の相関係数を求めることによって、CBMが個人内の特性を反映していることの信頼性を実証できるに違いない。そこで、No.1とNo.2のポイント総数の個人内の相関を求めたところ、相関係数は $r = 0.715$ ($t(42) = 6.62, p < 0.001$) で、有意に高かった。なお、No.1とNo.2の問題に対するポイント総数の散布図は、図4に示す通りである。

5) 問題の解答順序による影響

結果を分析する中で、最初に解いた問題と2番目に

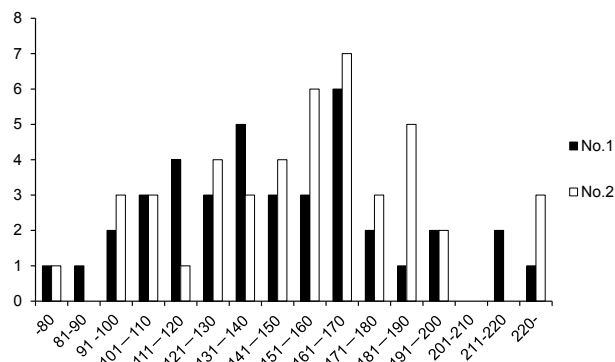


図3. No.1とNo.2のポイント総数による度数分布

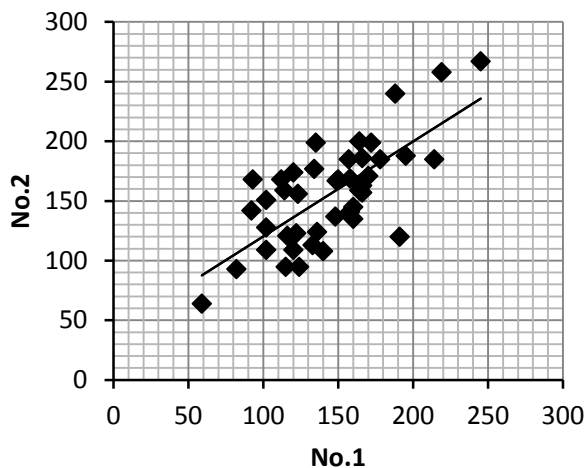


図4. No.1とNo.2のポイント総数による散布図

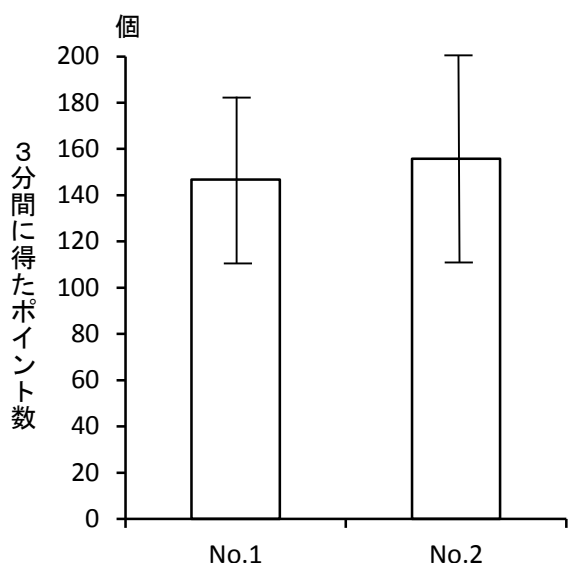


図2. ポイント総数によるNo.1とNo.2の問題の違い

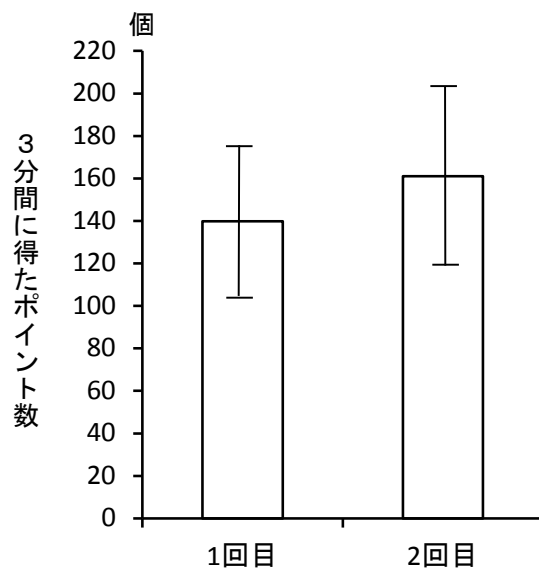


図5. 1回目と2回目の問題の提示順序によるポイント総数の違い

解いた問題によってポイント総数に差があることが視察された。そこで No.1 と No.2 の問題を併せて、1 回目に行った問題と 2 回目に行った問題とでポイント総数をまとめたものを図 5 に示す。問題の順番によって比較したところ、1 回目よりも 2 回目の方が得点有意に高いことが示された ($t(43) = 5.73, p < 0.001$)。

4. 考察

本研究の結果から、ポイント総数によって難しい問題と易しい問題として評価したものを比較したときに、図 1 に示すように正答数においても、またポイント総数においても No.1 に比べて No.2 の問題が有意に得点が高いことが示された。このことは、当該学年の問題を 5 割にし、残りを下学年のものにするというルールが同じでも、問題の難易度が異なってしまう可能性があることを示すものである。したがって、問題の難易度をそろえるためには、ポイント総数をほぼ近いものに設定する必要がある。No.1 の問題と No.2 の問題を比較したときに、No.1 の問題は最初の A4 版 1 枚の 18 問のポイント数が 144、次の 18 問のポイント数が 148 と問題の前半にポイントの高い問題続いていたのに対して、No.2 の問題は最初の 18 問のポイントが 110、次の 18 問のポイントが 142 ポイントであったことから、No.2 の問題の方が取り組みやすかったと推測される。

また、図 3 に示すように、ポイント総数による度数分布は正規性が棄却されなかったことから、正規分布していると考えられ、パーセンタイル順位によって被験者個人の相対的な位置を把握することができることから、妥当性をもつと考えられる。さらに、個人内の

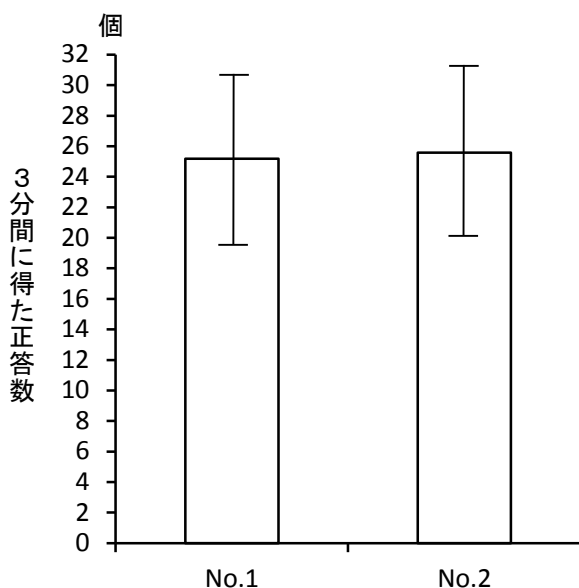


図 6. 正答数による No.1 と No.2 の問題の違い (研究 2)

No.1 と No.2 の問題における相関も有意であったことから、米国でのポイント総数による採点方法は信頼性があると判断することができる。

さらに、本研究の中で予想外だったのは、提示順によって問題のポイント数に有意な差があったことである。被験者は、本研究で初めて CBM の問題に答えた。したがって、第 1 回目のときには慣れていなかったものが、2 回目になると慣れてきたことによって、この有意差が出たと考えられる。

Ⅲ. 研究 2

1. 目的

研究 1 の結果から、ポイント総数による違いが難易度を表していることが考えられた。そこで研究 2 では、問題のポイント総数を同じにしたときに、被験者の解答した正答数やポイント総数に差が出るかどうかについて検討することにした。A4 版で作成した 18 問の問題の計 4 枚を、A3 版に両面印刷で 72 問とし 1 枚に収まるようにした。したがって、少なくとも A4 版 1 枚あたりのポイント総数が同じになるように設定した。仮説として、ポイント総数をそろえることによって、問題間の差は見られないであろうと予想した。

2. 方法

1) 被験者

被験者は、研究 1 と同じ大学生と専攻科学生 38 人 (女性 31 人男性 7 人) であった。

2) 算数の問題

ポイント総数が同じになるように設定した 2 つの間

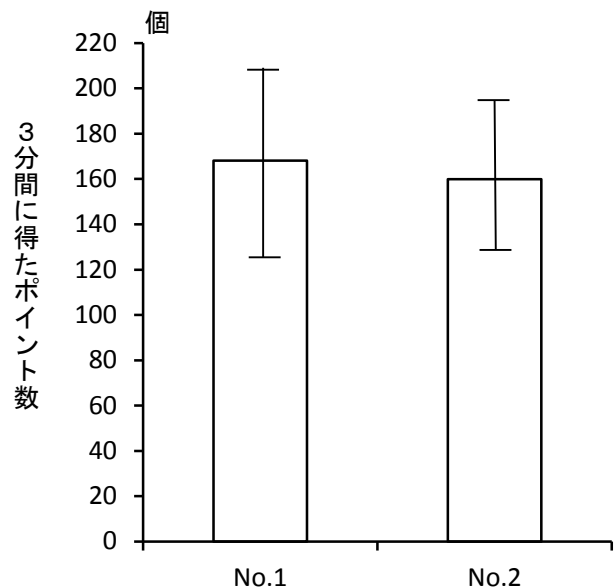


図 7. ポイント総数による No.1 と No.2 の問題の違い (研究 2)

題を用いた。No.1 の問題はそれぞれ 18 問 (A4 版 1 枚) のポイント総数が 132, 143, 131, 132 の計 538 であった。一方, No.2 の問題は, それぞれ 18 問のポイント総数が, 137, 140, 129, 133 の計 539 であった。

3) 手続き

手続きは, 研究 1 と同じであった。評価方法も研究 1 と同じであり, 提示順序による違いについても検討することにした。

3. 結果

1) 問題の正答数による比較

正答数による問題の比較として, 問題の平均正答数と ±1 標準偏差は, 図 6 に示す通りである。t 検定を行った結果, 両者の間には有意な違いは示されなかった ($t(37) = 0.58, ns$)。

2) ポイント総数による比較

米国式のポイント総数による評価による No.1 と No.2 の結果は, 図 7 に示す通りであった。図 7 に示すように, No.1 の問題と No.2 の問題には有意な差はなかった ($t(37) = 1.65, p > 0.10$)。

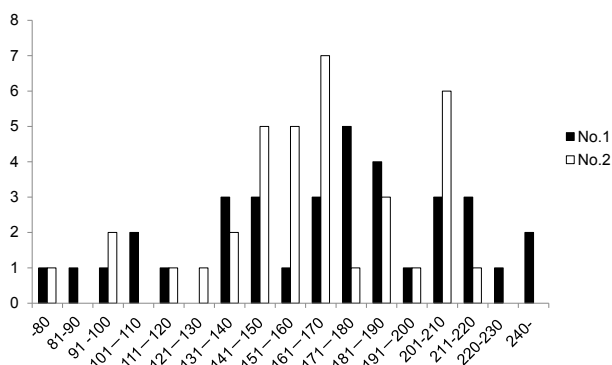


図 8. No.1 と No.2 のポイント総数による度数分布 (研究 2)

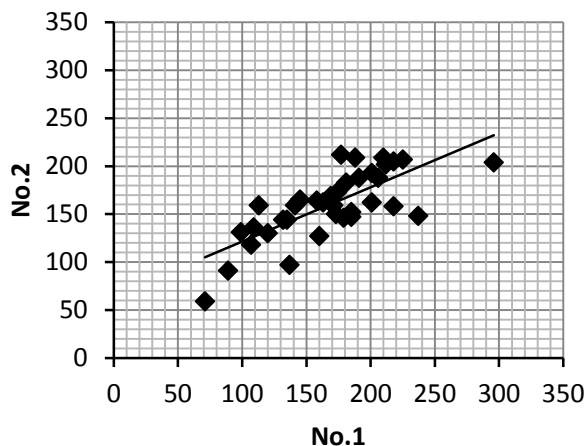


図 9. No.1 と No.2 のポイント総数による散布図 (研究 2)

3) 分布の正規性の検討

分布の正規性について, 研究 2 でも No.1 と No.2 それぞれの問題の正答数とポイント総数について尖度と歪度を求めたところ, いずれも有意ではなく, 正規性は棄却できなかった。ポイント総数による度数分布については, 図 8 に示す通りである。

4) No.1 と No.2 のポイント総数の個人内の相関

被験者は, 5 分の休憩を挟んで No.1 の問題と No.2 の問題を実施した。No.1 と No.2 のポイント総数の個人内の相関を求めたところ, 相関係数は, $r = 0.745$ ($t(36) = 6.70, p < 0.001$) で, 有意に高かった。なお, No.1 と No.2 の問題に対するポイント総数の散布図は, 図 9 に示す通りである。

5) 問題の解答順序による影響

研究 1 と同様に研究 2 でも, ポイント総数において最初に解いた問題と 2 番目に解いた問題によって差がありそうなことが視察された。そこで No.1 と No.2 の問題を併せて, 1 回目に行った問題と 2 回目に行った問題とでまとめたものを図 10 に示す。問題の順番によって比較したところ, 有意な差にまでは至らなかった ($t(37) = 2.66, p > 0.10$)。

4. 考察

ポイント総数をそろえることで二つの問題の間に差がないであろうと仮説を立てて実施した結果, 仮説通りに二つの問題の間には正答数とポイント総数の両方に有意差はなかった。このことからポイント総数をそろえることで, 問題間の差をなくすことができると結論づけることができる。

また, 研究 1 と同様に, 解答の順番によって正答数が異なるかどうか検討したが, 統計的には有意水準に

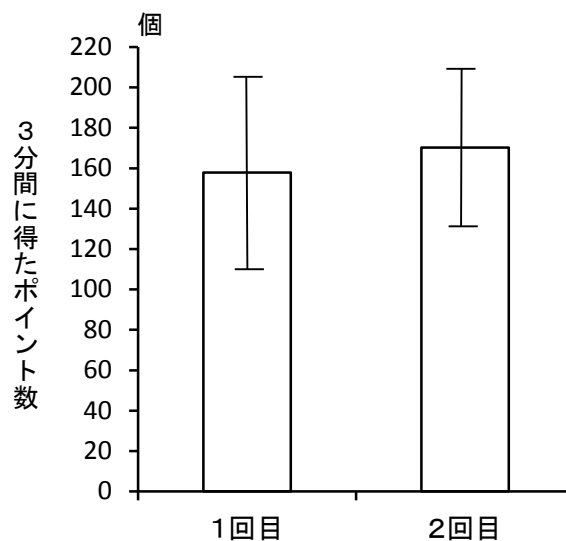


図 10. 1 回目と 2 回目の問題の提示順によるポイント総数の違い (研究 2)

は達しなかった。したがって、研究1では初めてCBMを実施したために要領を得ず1回目の問題での得点が下がったが、2回目の問題では要領がつかめたためにポイント総数が伸びたと考えられる。研究2は研究1よりも約2か月後に実施した。研究1と研究2の得点を比較すると研究2の得点が高かったという事実に基づけば、研究1での1回目と2回目の差はテストへの慣れが大きな要因であったと推測することができる。

IV. 総合考察

本研究の結果から、米国による採点方法を用いることで問題のポイント総数をそろえることができれば、CBMの問題間の難易度を統一することができることが明らかとなった。問題の等質性を保つことができれば、今後はいよいよ通常の学級の児童生徒を対象に、多人数の母集団でCBMを実施することとなる。多人数でのCBMの実施ができれば、それに基づいて対象の児童生徒の相対的なレベルを把握することができる。また、CBMを用いて定期的に進捗状況をモニターすることができれば、その成長比によって指導の適切性を判断することができるであろう。

Fuchs and Fuchs (1998) は、①特殊教育にかかる増加するコストの問題と、②マイノリティの児童の特殊教育の過剰出現の問題などのこれまでのディスクレパンシーモデルの問題点を明らかにし、CBMを用いた新しい認定方法としての4つのフェイズからなる適切性アセスメントプロセスを提唱している。この方法の基本的な考え方は、①すべての児童のための通常教育の可能性を最大にするように試みる、②通常教育での個々の児童生徒への調整の効果が検討するまで特殊教育のニーズに関する判断を保留する、③措置の前に学習を高める特殊教育プログラムを確かめるなど特殊教育に照会する前の通常教育を高めようとする、ことであった。Fuchs and Fuchs (1998) の4つのフェイズは、RTIと重なるところがあり、フェイズIでは通常の学級が十分な教育的環境であるかを評価、フェイズIIでは十分な教育環境の中での対象児童の二重の乖離を評価、フェイズIIIでは通常の学級内での調整が十分に行われているかを評価し、必要な学習状況に移行するかを評価、フェイズIVは特殊教育の有効性の評価となる。

なおフェイズIIにおいて評価される二重の乖離とは、CBMを用いて何週間か(7から10個)のデータポイントによってモニターされた進捗状況により、その子の同じ学級内での相対的なレベルと、個人内の成長比の二つから判断することになる。Fuchs and Fuchs (1998) は、ディスクレパンシーモデルに代わる学習

障害の児童生徒を認定する方法として、CBMによる進捗状況のモニタリングの結果としての二重の乖離から、学習障害の児童を特定する方法を提唱している。Fuchs and Fuchs (1998) は例として、レベルと成長比の二つの標準偏差の差からそれぞれ1.5標準偏差よりも低いもの(1.5/1.5基準)、レベルが1.5低い成長比は1.0標準偏差低いもの(1.5/1.0基準)、両方とも1.0標準偏差低いもの(1.0/1.0基準)を紹介している。またFuchs (2003) は、レベルのみ、傾き(成長比)のみ、レベルと成長比の二重の乖離で209人の母集団で反応のない児童を比較したときに、レベルのみで16人、傾きのみで24人、二重の乖離で9人であり、二重の乖離での値が学習障害の出現率と近かったことから、二重の乖離による評価が妥当であると述べた。Fuchs, Fuchs, and Speece (2002) は、二重の乖離モデルを用いた学習障害の特殊教育への認定の統一した構成概念として、介入に対する反応を二重乖離モデルの観点からとらえた処遇妥当性モデル(Treatment Validity Model)を提唱している。

したがってCBMは、学習障害児の認定に際してディスクレパンシーモデルに代わるアセスメントを可能にする重要な手立てであろう。わが国においてもCBMが標準化され、CBMに基づくレベルと成長の二重の乖離の方法が定着できれば、医療機関での診断に頼らずに教育的介入への応答(RTI)として、学習障害かどうかを判断することが可能になるであろう。本研究で得られた知見は、わが国でのCBMの標準化に向けた基礎的資料を提供するであろう。

米国のCBMに関する文献を展望する中で、コンピュータを活用したCBMが多く報告されるようになってきた。特に算数においては、計算問題を実施することで結果をグラフ化し、18までの被減数や繰り上がりの有無、2桁、3桁などの問題のパターンなどのスキル別の習熟度を示すものが開発されるようになってきた(例えば、Fuchs, Fuchs, Hamlett, & Stecker, 1990)。また読みの評価では、声に出して読む課題では一人一人の評価者が必要なことからMAZE課題を用いることで児童生徒の読みの進捗状況をコンピュータによってモニターすることができるようになってきた(Deno et al. 2008)。そこで今後のCBMのわが国における標準化に当たっては、計算やMAZE課題のようにコンピュータを活用したCBMの標準化が必要であろう。

参考文献

- Deno, S.L. (1985) Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52, 219-

- 232.
- Deno, S.L., & Fuchs, L.S. (1987) Developing curriculum-based measurement systems for data-based special education problem solving. *Focus on Exceptional Children*, **19** (8), 1-16.
- Deno, S.L., Reschly, A., Lembke, E., Magnusson, D., Callender, S.A., Windram, H., & Stachel, N. (2008) Developing school-wide progress-monitoring system. *Psychology in the Schools*, **46** (2), 44-55.
- Fuchs, D., Mock, D., Morgan, P.L., & Young C.L. (2003) Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice*, **18** (3), 157-171.
- Fuchs, D., Roberts, P.H., Fuchs, L.S., and Bowers, J. (1996) Reintegrating students with learning disabilities into the mainstream: A two-year study. *Learning Disabilities Research and Practice*, **11** (4), 214-229.
- Fuchs, L.S. (2003) Assessing intervention responsiveness: Conceptual and technical issues. *Learning Disabilities Research and Practice*, **18** (3), 172-186.
- Fuchs, L.S., & Fuchs, D. (1998) Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research and Practice*, **13** (4), 204-219.
- Fuchs, L.S., Fuchs, D., & Fernstrom, P. (1993) A conservative approach to special education reform: Mainstreaming through transenvironmental programming and curriculum-based measurement. *American Educational Research Journal*, **30** (1), 149-177.
- Fuchs, L.S., Fuchs, D., Hamlett, C.L., & Stecker, P.M. (1990) The role of skills analysis in curriculum-based measurement in math. *School Psychology Review*, **19** (1), 6-22.
- Fuchs, L.S., Fuchs, D., & Speece, D.L. (2002) Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly*, **25**, 33-45.
- Gilbert, J.K., Compton, D.L., Fuchs, D., Fuchs, L. S., Bouton B., Barquero, L. A., & Cho, E. (2013) Efficacy of first-grade responsiveness-to-intervention prevention model for struggling readers. *Reading Research Quarterly*, **48** (2), 135-154.
- Grigorenko E.L. (2009) Dynamic assessment and response to intervention: Two sides of one coin. *Journal of Learning Disabilities*, **42** (2), 111-132.
- 廣瀬由美子・桂 聖・坪田耕三 (2009) 通常の学級担任がつくる授業のユニバーサルデザイン. 東洋館出版社.
- 干川 隆 (1999) アメリカ合衆国における学習障害の認定をめぐる問題－能力と学力の不一致モデルの検討－ 国立特殊教育総合研究所研究紀要, **26**, 97-110.
- 干川 隆 (2000) 学習障害児と低学力児との類似性をめぐる論争－米国ミネソタ大学グループの研究成果の照会－ 国立特殊教育総合研究所研究紀要, **27**, 89-99.
- Individual with Disabilities Education Improvement Act. 20 U.S.C. § 1400 *et seq.* (2004)
- Jenkins, J.R. & Pany, D. (1978) Standardized achievement tests: How useful for special education? *Exceptional Children*, **44**, 448-453.
- 海津亜希子 (2006) 日本における LD 研究への示唆－米国での LD 判定にみられる変化をうけて－ LD 研究, **15** (2), 225-233.
- 海津亜希子・田沼実敏・平木こゆみ・伊藤由美・Vaughn S. (2008) 通常の学級における多層指導モデル (MIM) の効果－小学校 1 年生に対する特殊音節表記の読み書き指導を通じて－ 教育心理学研究, **56**, 534-547.
- Kratochwill, T.R., Volpiansky, P. Clements, M., & Ball, C. (2007) Professional development in implementing and sustaining multitier prevention models: Implications for response to intervention. *School Psychology Review*, **36** (4), 618-631.
- Marston, D. (1988) The effectiveness of special education: A time series analysis of reading performance in regular and special education setting. *The Journal of Special Education*, **21** (4), 13-26.
- Marston, D. (1989) A curriculum-based measurement approach to assessing academic performance: What it is and why do it. Shinn M.R. (ed.) *Curriculum-Based Measurement: Assessing Special Children*, NY: The Guilford Press, 18-78.
- 美坂昌宏 (2006) 学習につまずきのある児童への学習支援と評価方法に関する研究. 熊本大学大学院教育学研究科修士論文
- Salmon-Cox, L. (1981) Teachers and standardized achievement tests: What's really happening? *Phi Delta Kappan*, **62** (9), 631-634.
- Shapiro, E.S., & Derr, T.F. (1987) An examination of overlap between reading curricula and standardized achievement tests. *The Journal of Special Education*, **21** (2), 59-67.
- Tindal, G.A., & Marston, D.B. (1990) Math assessment. G.A. Tindal and D.B. Marston, *Classroom-Based Assessment: Evaluating Instructional Outcomes*, Merrill Publication Company.
- Torgesen J.K. (2002) The prevention of reading difficulties. *Journal of School Psychology*, **40** (1), 7-26.
- U.S. Department of Education (2002) *A New Era: Revitalizing Special Education for Children and Their Families*. President's Commission on Excellence in Special Education.