

# 認知特性の評価に使用される米国版検査の文化負荷量と言語要求度

島田 直子・今村 由未香

Cultural loading and Linguistic demand in cognitive measurements used in the United States

Naoko Shimada (Department of Psychology, Komazawa University, Japan)

Yumika Imamura (Kawagoe Juvenile Prison, Japan)

## 要旨

近年、多様な文化と言語の影響を受ける外国ルーツをもつ子どもたちが増えている。しかしながら、そのような子どもたちの心理教育的アセスメントに関する国内の研究は非常に限られる。多様な文化や言語の背景のある子どもが数多く教育を受ける米国では、認知特性の評価に使われる検査における文化と言語の影響について、文化と言語の解釈表 (Culture-Language Interpretive Matrix: C-LIM) が開発された。本稿では我が国における C-LIM の利用可能性を検討するために C-LIM の概要についてまとめ、日本で使用されている主要な検査のうち米国に開発の背景がある検査に焦点をあて、それぞれの開発の元になった米国版検査の C-LIM を参考に下位検査の文化負荷量と言語要求度について記した。最近の研究では C-LIM の臨床利用の限界について示唆している論文が多くある。その点も踏まえて日本のアセスメント研究において C-LIM 関連の研究が役立つ可能性について論じた。

The current study reviewed literature about studies related to Culture-Language Interpretive Matrix: C-LIM for psychoeducational assessment of children from culturally and linguistically diverse backgrounds. This literature review revealed that the validity and clinical utility of the C-LIM was not warranted; however, the C-LIM could help researchers and practitioners in that it posited a framework of understanding subtle differences in terms of their cultural loading and linguistic demand. Future research should be noted for the clinical application of the C-LIM.

**KEY WORDS:** 外国人児童生徒, 認知特性の評価, 知能検査, 心理教育的アセスメント, Culturally and linguistically diverse children, Cognitive assessment, Intelligence tests, Psychoeducational assessment

## 【問題と目的】

近年、多様な文化的背景のある子どもが増加している。2020 年末現在、日本の在留外国人数は 276 万 635 人であると報告されている (法務省入国管理局, 2021)。また、日本国籍であっても日本語の指導を必要としている子どもや文化的背景が日本とは異なる子どももいる。このような外国にルーツをもつ子どもは、文化や言語の違いから、学習や学校生活への適応で支援を必要とする場合がある (文部科学省, 2019)。支援に際して、子どものニーズを正確に見立てることが必要になるが、実際には、日本とは文化や言語が異なることに起因すると考えられる学校での困難さが、障害に起因する判断されている可能性が指摘されている (e.g. 金, 2020; 松田・中川, 2017)。

このようなケースが起こっているのはわが国だけではない (e.g. Herrera, 2020; Rhodes et al.,

2005)。米国の研究では、この問題の背景には様々な要因があると考えられているが、支援の現場での課題に注目すると、通常学級における指導で文化と言語の多様性に配慮した指導が行われているか、アセスメントに至る過程で子どもの背景の多様性への理解に課題はないか、心理アセスメントにおいて適切かつ十分な情報が考慮されているかなどの視点があげられる (Rhodes et al., 2005)。わが国の外国にルーツをもつ子どもの支援においては、適切な指導環境の整備とともに、子どものニーズに応じた支援につなげるためのアセスメント方法の検討が急務である。

発達障害等の判定を含む子どもの心理教育的アセスメントにおいては認知特性の評価が必要になる場合があるが、文化や言語的背景の異なる外国にルーツのある子どもに通常の日本版の知能検査を実施することには課題がある (島田, 2016)。日本版の検査内容を子どもの母語に翻訳する (緩利

ら, 2010), 検査実施に通訳を利用する (丸山, 2013), 図形などを使用したマトリクス検査や母語版検査を利用する (松田ら, 2016; 二井ら, 2013) など報告があるものの, 認知特性の評価についての国内の研究は非常に限られる。

鳥田 (2016) は米国における多文化アセスメントに関する文献のレビューを通して, 多文化や多言語の背景のある子どもへの知能検査の利用方法を整理して報告している。その中には, 子どもの状況に合わせた実施手順の変更, 通訳の利用, 母国版検査や非言語性検査の利用, 文化と言語の解釈表 (Culture-Language Interpretive Matrix: C-LIM) を活用した得点の妥当性の確認, バイリンガルの言語発達を考慮した検査の選択などが含まれる。各方法には長短があり逸出する方法がないと (Flanagan et al., 2013) されている中で, わが国でのアセスメントの可能性について論じている。その一つに C-LIM の利用について触れられているが, 概要が端的に述べられているのみであるため, 実用性について検討するための情報が不足している。そこで本稿では, この C-LIM に関するより詳細な情報を整理して報告することを目的とする。

## 【方法】

はじめに, 鳥田 (2016) に引用されている C-LIM の関連文献をレビューし, C-LIM の概要について, 開発の背景, 特徴, 利用方法についてまとめた。次に, 日本で使用されている主要な検査のうち, 米国に開発の背景がある検査に注目し, 各検査の原版である米国版の C-LIM を参考に検査の特徴を整理した。加えて, 各検査の検査マニュアル, またはそれに準ずる文献から, C-LIM 利用に関する見解が示唆される記述について記した。最後に, C-LIM というキーワードを用いて, EBSCOhost にて文献検索を行った。Academic Search Premier, APA PsychArticles, MEDLINE, APA PsycInfo, Psychology and Behavioral Sciences Collection を選択し, 対象期間は, 1974 年～2022 年までとした。検索の結果, 319 件の文献がヒットした。論文のタイトルから重複する文献と関連がないと判断できる文献を除外した。その結果 7 本の学術論文がヒットした。その学術論文を取り上げ C-LIM に関する研究の動向をまとめた。

## 【結果】

上記の方法で文献レビューを行った結果について, 1. C-LIM とその利用法, 2. 日本版が開発されている米国版検査の C-LIM, 3. C-LIM 利用に関する最新の研究動向の順に報告する。

### 1. 文化と言語の解釈表(C-LIM)とその利用法

C-LIM の概要 (Flanagan & Alfonso, 2011) については邦訳 (上野・名越, 2013) があり, 文化 - 言語解釈マトリクスとも訳されている。C-LIM とは, 認知特性の評価に使用される各種検査について, 文化的知識や経験が必要とされる程度を示す文化負荷量と, 検査で使用される言語の力が要求される程度を示す言語要求度を, 下位検査ごとに評価したものである。文化負荷量は, 認知の過程に関する検査か, その産物物に関するものかという点, 検査刺激が抽象的で新奇性のあるものか, 特定の文化と関連の強いものかという点, 言語による教示か, ジェスチャーなどが使われるかという点に注目している。言語負荷量については, 受容言語と表出言語のそれぞれの要求具合に加え, 教示に言語が必要か, その言語表現が長文で高度な理解力を必要とするか, 絵やジェスチャーなどが用いられる理解が容易なものかが考慮される (Flanagan et al., 2007)。

C-LIM では, 先行研究における多文化や多言語の背景のある子どもの検査得点 (e.g., Cummins, 1984) を参考に, 専門家の意見 (Expert opinion) をもとにして, 文化負荷量と言語要求度が, それぞれ, 大, 中, 小, の 3 段階で評価されている。横軸に言語要求度, 縦軸に文化負荷量がそれぞれ  $3 \times 3$  (e.g. 大, 中, 小) で配置され, その組み合わせ (e.g., 言語要求度低  $\times$  文化負荷量大) が 9 つのセルに整理される。各下位検査は, それぞれの文化負荷量と要求度の大きさに応じて, その 9 つのうちのいずれかに分類されている (Flanagan et al., 2013; Ortiz, 2017)。

このような検査側の特徴に加え, 受検者側の特徴も考慮される。国外に文化や言語の背景がある子どもたちの中でも, 子どもの居住国言語の習熟度や文化的な知識や経験量には個人差がある。そこで, 目安として居住年数, 文化適応度, 言語習得度などが考慮され, 米国の家庭で生まれ育った子どもたちとの差の大きさが, 大, 中, 小の 3 段

階で示される。差が小さい場合の例としては、米国在住歴が概ね7年以上で文化適応もかなり進んでおり、渡米前の教育環境も安定している。ネイティブ並みではないが、移住後2年から3年で習得されるとされる日常会話に必要な基本的対人伝達能力(Basic interpersonal communicative skills; BICS)は高度に発達し、学習に必要な認知学習的言語能力(Cognitive and academic language proficiency; CALP)の発達もみられる。両親に高校卒業以上の教育歴があり、米国ネイティブ並みの会話力と読み書きの力が安定している。中程度の場合は、米国在住3年から7年程度で文化適応も中程度である場合が多い。本人に米国の現地語である英語での会話力はあるが、両親の教育歴が比較的短く英語があまり話せないケースなどがある。差が大きい場合は、BICSのレベルも文化適応の度合いも低く、滞在歴は3年以下である場合が多い。入国前の教育環境が安定していない場合や英語での会話力も読み書き力も学習し始めたばかりという場合がある(Flanagan et al., 2013)。C-LIMでは、受検者と米国の子どもの差の大きさに応じて、各下位検査得点で予想される低下度が3段階(e.g., 5~7点, 7~10点, 10~15点)で表示されている。この得点は、平均100, 標準偏差15とした場合である。

## 2. C-LIMの利用の方法

当初、C-LIMは主訴に応じて文化と言語の違いによる影響が少ない検査を選択するための評価表の開発を目的としていた。しかし、Gv(視覚的知能)、Gsm(短期記憶)、Gf(流動性知能/推論)(Flanagan et al., 2007; 三好・服部, 2010)に関する下位検査を除いて、文化負荷量と言語要求度がともに低いものはほとんどないということが研究の初期段階で明らかになったため、視座の転換を図り文化と言語の影響による得点の低下度を臨床場面で得られた得点の解釈に利用するという方法として開発された(Flanagan et al., 2007)。なお、検査選択の際に参考にすべきとの見解もある(Cormier, 2014)。前述のようにC-LIMの開発の方法が専門家の判断をもとにしていることから、C-LIMの妥当性に関するエビデンスが不十分であるという指摘もある。C-LIMの利用に関しては多くのワークショップなどが行われているが、その妥当性の検証はほとんどが博士論文としての

研究であるとの批判もある(Styck & Watkins 2013)。

C-LIMの利用方法(Ortiz, 2017; Ortiz & Wong, 2020)について概略すると、まず、通訳などは使用せず、子どもの居住国である米国版の検査を標準実施する。文化負荷量と言語要求度が大きいとされる検査の得点が低く、その程度が小さいとされる検査の得点が高い傾向が見られれば、検査得点が文化と言語の違いによる影響を受けている可能性が高いと判断される。C-LIMの想定以上の得点低下が見られた下位検査については、母語版検査か通訳を通して米国版検査を再実施し得点の低下度合いを確認する。バイリンガル検査者も通訳も利用できない場合、妥当性は下がるが、課題観察などのインフォーマルな方法で補う(Ortiz & Wong, 2020)。Gc(結晶性知能)に関しては、言語と文化的知識そのものを測定する検査であるため、文化負荷量と言語要求度を検査の本質と切り離して考えることが難しい。そこで、同じような発達背景の子どもと比較することを目的としてPicture Vocabulary Acquisition Test(PVAT: Ortiz)という英語の受容言語検査を利用する方法も示されている。PVATは米国内の様々な言語背景の子どもについて英語の学習経験年数ごとに標準化された検査で、母語背景の違いによる得点の有意差はないとされている(Ortiz & Wong, 2020)。

米国版WISC-V関連の文献(Flanagan & Alfonso, 2017)によれば、米国では南米からの移民向けにスペイン語版WISC-V(Wechsler, 2017)が開発されているため、母語版検査についてはその検査を利用することが出来るが、それ以外は子どもの母国の検査を利用する。米国国外で標準化された検査は、米国で育ち母国での教育を十分には受けていない子どもについては考慮していないことが想定される。そのような母語版検査を利用した場合の得点が、どのような傾向を示すかについてはほとんど研究がなされていない(Flanagan & Alfonso, 2017)。一方で、少なくとも米国版に関しては文化や言語マイノリティの子どもの得点の低下傾向が明らかにされている。このような背景から、米国版検査の得点評価が優先され、母語版検査はその結果の妥当性を向上させる目的で使われる。

母語版検査の利用はあくまでも補助的であるた

め、練習効果 (Practice Effect) については、それほど重要視していない (Flanagan & Alfonso, 2017)。練習効果によって、母語版検査の得点が平均レベルに達したとすれば、それは適切に学習をすれば積み上げができるというエビデンスともいえるし、再検査で平均範囲の得点が示されている場合に、認知的な障害があることは非常にまれであると述べられている (Flanagan & Alfonso, 2017)。ただし、使用言語が変わる場合の練習効果についてエビデンスが示されているわけではない。同一言語の場合の WISC-IV の FSIQ の練習効果に関しては、9~13 か月後の再実施得点が 1.63 点上昇 (n = 43)、標準化サンプルの 11 か月後の得点では 5.60 点の上昇と報告 (Ryan et al., 2010) があるが、バイリンガル検査者による再検査では、これ以上の得点上昇があるケースも報告されている (Diana v. State board of education, 1970; Jacob & Hartshorne, 2003)。

### 3. 米国版検査の C-LIM

米国では学校現場で心理教育的アセスメントを行うことが珍しくない。2010 年から 2011 年にかけて米国で実施された認知面のアセスメントに関する実態調査 (Sotelo-Dynega & Dixon, 2014) では、調査に参加した 323 名のスクールサイコロジスト (School psychologist; 以下 SP とする) の 9 割以上が、言語や文化の違いを考慮して、通常とは異なるアセスメント方法を採用すると回答している。具体的に利用する検査についての質問も含まれており、多文化背景の子どもに利用された報告のある米国版検査のうち、日本版が刊行されているものが 3 検査ある。日本版 WISC-IV 知能検査 (日本版 WISC-IV 刊行委員会, 2010a) の開発のもとになっている Wechsler Intelligence Scale for Children-Fourth Edition (Wechsler, 2004)、日本版 Kaufman Assessment Battery for Children Second Edition (KABC-II: 日本版 KABC-II 制作委員会, 2013a) の開発のもとになっている KABC-II: Kaufman Assessment Battery for Children, Second Edition (Kaufman & Kaufman, 2004a)、日本版 DN-CAS 認知評価システム (DN-CAS: 前川・中山・岡崎, 2007a) の開発のもとになっている Cognitive Assessment System (CAS: Naglieri & Das, 1997) である。

なお、米国の最新のビネー式検査である

Stanford-Binet Intelligence Scale-Fifth Edition (SB5; Roid, 2003) についても報告があったが、日本で現在利用されている田中ビネー知能検査 V (田中ビネー V: 田中教育研究所, 2005) は、1937 年のスタンフォード版を基礎とし、1947 年の初版から改訂を重ね日本の文化や生活様式等を反映して、より日本人に適した検査として発展している (杉原・杉原, 2005)。田中ビネー V は、Binet の知能観に基づいてはいるが、原版の検査を忠実に翻訳したものではない (杉原・杉原, 2005)。田中ビネー V と SB5 は検査構成も異なる。

日本版 KABC-II には認知尺度と習得尺度があるが、米国版では、認知尺度と習得尺度が独立している (Kaufman & Kaufman, 2004b)。習得尺度は学力検査であり、学力はある社会の成員として獲得が期待される基礎知識を指し知能とは異なる (荘島, 2014)。日本版の認知尺度に相当する米国版 KABC-II (Kaufman & Kaufman, 2004a) に含まれる 18 の下位検査のうち 11 の検査が日本版の認知尺度に採用されている (日本版 KABC-II 制作委員会, 2013b)。日本版の WISC-IV と DN-CAS の開発においては、下位検査の構成に変更はない (日本版 WISC-IV 刊行委員会, 2010b; 前川他, 2007b)。なお、日本版 WISC-IV については、改訂版 (日本版 WISC-V 刊行委員会, 2022a) が刊行されている。日本版 WISC-V については、米国版の 21 の下位検査のうち 5 つの関連指標の下位検査を除く 16 の下位検査から構成されている (日本版 WISC-V 刊行委員会, 2022b)。

また、日本の子どもに適した言語や文化を考慮した検査項目の採用 (日本版 KABC-II 制作委員会, 2013b; 日本版 WISC-IV 刊行委員会, 2010b; 日本版 WISC-IV 刊行委員会, 2022b)、難易度の調整 (前川他, 2007b)、天井効果改善等のための項目追加 (日本版 KABC-II 制作委員会, 2013b; 日本版 WISC-IV 刊行委員会, 2022b)、などが行われているが検査の本質に関わる変更はないようである。このことから、日本版検査についても米国版の文化負荷量と言語要求度の傾向は、ある程度類似すると仮定できるのではないかと考えられる。そこで、表 1 に米国版 WISC-IV と WISC-V、表 2 に KABC-II と CAS の C-LIM を示した。米国版検査については、KABC-II の再標準化版 (Kaufman & Kaufman, 2018b) と CAS の改訂版 CAS2 (Naglieri et al., 2017) が刊行されているが、

表 1 米国版 WISC-IV と WISC-V の下位検査の文化負荷量と言語要求度と日本語版下位検査の対応 (平均 100・標準偏差 15)

米国語版下位検査	米国版検査における受検者の文化・言語差の程度によって低下すると推定される得点										米国版の各下位検査に対応する 日本語版下位検査	
	小:3-5点 中:5-7点 大:7-10点	言語低 文化中	言語中 文化低	言語低 文化高	言語中 文化中	言語高 文化低	小:7-10点 中:10-15点 大:15-20点	言語中 文化中	言語高 文化中	言語高 文化高		
<b>WISC-IV</b> (Wechsler, 2004)												
Picture Completion												絵画完成
Vocabulary												単語
Coding												符号
Similarities												類似
Block Design/BD No time bonus												積木模様/積木模様・時間割増なし
Arithmetic												算数
Matrix Reasoning												行列推理
Digit Span												数唱
Information												知識
Comprehension												理解
Symbol Search												記号探し
Letter-Number Sequencing												語音一覧
Cancellation												絵の抹消
Picture Concepts												絵の概念
Word Reasoning												語の推理
<b>WISC-V</b>												
<b>(Wechsler, 2014)</b>												
Similarities												類似
Vocabulary												単語
Information												知識
Comprehension												理解
Block Design												積木模様
Visual Puzzles												パズル
Matrix Reasoning												行列推理
Figure Weight												バランス
Picture Concepts												絵の概念
Arithmetic												算数
Digit Span (Back word)												数唱(逆唱)
Picture Span												絵のスペン
Letter-Number Sequencing												語音一覧
Coding												符号
Symbol Search												記号探し
Cancellation												絵の抹消
Naming Speed Literacy												該当なし(呼称速度・文字)
Naming Speed Quantity												該当なし(呼称速度・数量)
Immediate Symbol Translation												該当なし(即時シンボル変換)
Delayed Symbol Translation												該当なし(遅延シンボル変換)
Recognition Symbol Translation												該当なし(再認シンボル変換)

表2 米国版 KABC-II と CAS の下位検査の文化負荷量と言語要求度と日本語版下位検査の対応 (平均 100・標準偏差 15)

米国語版下位検査	米国版検査における受検者の文化・言語差の程度によって低下すると推定される得点												米国の各下位検査に対応する 日本語版下位検査
	小:3-5点 中:5-7点 大:7-10点	言語低 文化低	言語中 文化中	言語低 文化高	言語中 文化中	言語中 文化中	言語高 文化低	言語中 文化高	言語中 文化中	言語高 文化中	言語高 文化中	言語高 文化高	
<b>KABC-II</b> <b>(Kaufman &amp; Kaufman, 2004a)</b>													<b>KABC-II</b> <b>(日本語版KABC-II制作委員会, 2013a)</b>
Atlantis	○												語の学習
Face Recognition	○												顔さかし
Story Completion								○					物語の完成
Number Recall													数唱
Gestalt Closure			○										絵の統合
Atlantis Delayed	○												語の学習→遅延
Rover					○								近道探し
Triangles	○												模様構成
Word Order						○							語の配列
Pattern Reasoning	○												パターン推理
Hand Movements	○												手の動作
Expressive Vocabulary													習得尺度検査(表現語彙)
Riddles													習得尺度検査(なぞなぞ)
Verbal Knowledge													習得尺度検査(理解語彙)
Block Counting					○								該当なし(積木さがし)
Rebus					○								該当なし(文の学習)
Rebus Delayed					○								該当なし(文の学習→遅延)
Conceptual Thinking								○					該当なし(仲間探し)
<b>CAS</b>													<b>DN-CAS</b>
<b>(Das &amp; Naglieri, 1997)</b>													<b>(前川・中川・岡崎, 2007a)</b>
Figure Memory	○												図形の記憶
Nonverbal Matrices	○												図形の推理
Expressive Attention	○												表出の制御
Receptive Attention			○										形と名前
Matching Numbers					○								数の対探し
Planned Codes					○								文字の変換
Number Detection					○								数字探し
Word Series							○						単語の記憶
Planned Connections								○					系列つなぎ
Verbal Spatial Relations									○				関係の理解
Sentence Repetition													文の記憶
Sentence Questions													統語の理解

本稿では下位検査の対応などを考慮し、日本版開発当時の原版を参照する。米国版検査の各下位検査について、日本版検査との対応について併記したが、これは読者に分かりやすく示すことを意図して表示したもので、各検査のC-LIMが日本版検査にも当てはまることを示しているわけではない。

米国版 WISC-IV の C-LIM (Ortiz, 2023; Ortiz & Wong, 2020) をもとに、下位検査の文化負荷量と言語要求度を表 1 に示した。米国版 WISC-IV の 15 の下位検査のうち、文化負荷量、言語要求度ともに低いのは「絵の抹消」と「行列推理」である。「理解」、「知識」、「類似」、「単語」、「語の推理」は文化負荷量、言語要求度ともに高く (Ortiz, 2023)、言語や知識に関わる課題は特に得点の低下度が大きい。「符号」、「積木模様」、「数唱」、「記号探し」など、文化の違いによる影響を受けにくいとされる検査でも得点の低下傾向が見られる。

米国版 WISC-V では、新しい下位検査として「パズル」「バランス」「絵のспан」が含まれている。WISC-V の C-LIM (Ortiz, 2023) によれば、「パズル」は文化負荷量、言語要求度ともに低い。「バランス」は文化負荷量、言語要求度ともに中程度、「絵のспан」は文化負荷量が低く、言語要求度は中程度とされている。米国版 WISC-IV では、文化負荷量、言語要求度ともに高かった「理解」は WISC-V では、言語要求度が高く、文化負荷量は中程度とされていた。また、米国版 WISC-IV で文化負荷量が低く、言語要求度が中程度であった「数唱」は、WISC-V では、文化負荷量が低く、言語要求度が高いという評価となっている (Ortiz, 2023)。

米国版 KABC-II の下位検査の文化負荷量と言語要求度については、Ortiz (2023) に示されている C-LIM を参考にした。日本版の開発では、米国版 KABC-II の 18 の下位検査のうち、「積木さがし」、「仲間探し」、「文の学習」、「文の学習-遅延」の 4 検査が削除され、「表現語彙」、「なぞなぞ」、「理解語彙」の 3 検査は習得尺度に採用されている (日本版 KABC-II 制作委員会, 2013b)。認知尺度として採用された 11 の下位検査のうち 6 検査が、文化負荷量、言語要求度ともに低くなっている (Ortiz, 2023)。

CAS の C-LIM は Flanagan et al. (2007) に記載があったが、同書の改訂版 (Flanagan et al.,

2013) には CAS の C-LIM に関する言及はない。米国では Cognitive Assessment System-Second Edition (CAS2; Naglieri et al., 2014) が開発されたため、最新の C-LIM (Ortiz, 2023) には CAS2 の C-LIM が掲載されているが、CAS については掲載がない。そのため、CAS の C-LIM については Flanagan et al. (2007) 版を参考にする。CAS は「図形の記憶」、「図形の推理」、「表出の制御」の 3 つの下位検査が文化負荷量、言語要求度ともに低くなっている。文化負荷量、言語要求度がともに高い下位検査は含まれていない。「文の記憶」と「統語の理解」が、CAS の中で一番得点の低下度の大きい下位検査となっており、言語要求度が高く、文化負荷量が中程度となっている。なお、5~7 歳の子どもに実施する「発語の速さ」については C-LIM の記載がなかった (Flanagan et al., 2007)。

各検査関連のマニュアル等で、C-LIM の利用がどのように位置付けられているかについて情報収集を試みた。米国版 WISC-V の実施マニュアル (Wechsler, 2014) には、WISC-V は英語を話す米国人を対象としており、米国在住の英語を第一言語とする子どもの得点に基づいて標準化されている点に留意し、通訳や翻訳の利用など、定められた手続きから外れた方法で実施する場合には、その解釈に特に注意して、臨床的判断をする必要があるとの記載がある。米国版 WISC-V では、言語が変わっても認知的障害は変化しないという交差言語的確認 (Cross linguistic confirmation) を前提として、C-LIM の利用も示されている (Flanagan & Alfonso, 2017)。

DN-CAS の米国改訂版である CAS2 では、通訳の利用などと並んで、C-LIM の利用にも触れられているが、その具体的な方法が示されているわけではなく、C-LIM の実用には課題があることが論じられている (Naglieri & Otero, 2017)。C-LIM は、検査の文化的・言語的影響について理解が不足している場合に、得点の傾向を参考にできる可能性はあるが、専門家の意見を基に作成されたもので、診断や特別支援教育の必要性を査定するためのツールとしては、十分に信頼性が確立されているとは言えないとしている (Naglieri & Otero, 2017)。米国版 KABC-II については、2018 年に再標準化が行われたが、本稿で参考した関連文献 (Kaufman et al., 2005) は改訂されておらず、

米国版 KABC-II の多文化への適用における C-LIM の位置づけを本稿で確認することは困難であった。

#### 4. C-LIM の妥当性に関する研究

Cormier et al. (2014) は、Woodcock-Johnson Tests of Cognitive Abilities - Third Edition (WJ III) の再標準化データ (n=8,782) を使用し、文化負荷量、言語要求度、認知検査でのパフォーマンスの3つを潜在変数とした構造方程式モデルを用いてモデルの適合度の確認を行った。文化負荷量については、人種、家庭内言語、第一言語、国外生まれかどうかを観測変数とし、言語要求度については、WJ III の学力検査 (WJ III ACH) から4つの口語検査得点が観測変数として採用された。また、WJ 認知尺度 (WJ III COG) の得点が認知検査でのパフォーマンスの観測変数とされた。言語要求度の認知検査のパフォーマンスへの影響については比較的パス係数が高かったが、文化負荷量についてはあまり高くなく、使用された観測変数が適切ではない可能性が考えられた。Cormier et al. (2014) は文化負荷量のより適切な観測変数が明確にされる必要があると述べている。

Cormier et al. (2016) では、下位検査の言語要求度という観点から WISC-IV と WISC-V の口語教示に注目し、教示文の冗長さ (単語数、文章数、3音節以上の単語数) と複雑さ (1つの文章に使われている単語数の平均、1つの単語の中の音節数の平均) を数値化し下位検査間の比較を行った。C-LIM で予測される言語要求度には検査間にばらつきがあるが、教示の特徴のみに注目すると下位検査間の言語要求度にはそれほど大きな差はなかった。この結果は、下位検査の言語要求度については、口語教示以外の側面にも注目することの重要性を示している (Cormier et al., 2016)。また、WISC-V では WISC-IV に比べて教示が平易化されているという開発元の主張と一致する傾向が見られた下位検査 (e.g., 積木模様、語音整列、理解) があると報告している (Cormier et al., 2016)。

Hasson et al. (2019) は「積木模様」、「行列推理」、「単語」、「類似」の4つの下位検査から構成されるウェクスラー式簡易検査 WASI-II に関して、在米期間が3年以上で英語を流暢に話することができる12~17歳のアラブ系アメリカ人男子の得点 (n=60) が C-LIM の予測と一致するかを検証し

た。「積木模様」と「行列推理」の得点はほぼ C-LIM と一致したが、「単語」と「類似」については C-LIM よりも高くなった。また、「類似」に比べて「単語」の得点傾向が高く、英語学習が進むことで「単語」の得点は高くなるが、複数の異なる対象に共通する要素を抽出する際の抽象化の度合いや分類形態は文化の影響を受けやすいため「類似」の得点の方が低くなったのではないかと論じている。さらに Hasson et al. (2019) は文化適応度の検査得点への影響に注目し、文化適応尺度の得点を説明変数に含め階層的重回帰分析を行ったが、両親の収入や母親の教育歴等の社会経済的背景を統制すると、文化適応度の認知検査得点への優位な影響については確認できず、この結果からも文化負荷量の推定の難しさが示唆された。

Kranzler et al. (2010) と Calderon-tena et al. (2020) は WJ III COG の得点を用いて C-LIM の有用性と妥当性について検証した。Kranzler et al. (2010) の研究では、5~18歳の第二言語としての英語教育を受ける子どもの得点 (n=46) について、その得点がどの程度 C-LIM の示す得点傾向に一致するかについて検証した。全体の平均としては、概ね C-LIM の示す傾向が見られたものの、個々の子どもの得点についての検証では C-LIM の示す得点の低下パターンに13%の子どもの得点しか一致しなかった。前述したように C-LIM では受検者の第二言語の習得度等を考慮して得点の低下度合いを3段階で予測するが、41%の子どもたちはどのパターンにも当てはまらないという結果になった。

Calderon-tena et al. (2020) の研究では、母語が英語ではない在米期間の短い子ども (n=46) と、第一言語が英語以外の言語で完全に米国文化に適応しているわけではないという点で文化的マジョリティの子どもと異なるが、最初のグループの子どもよりは適応が進んでいると考えられる子ども (n=147) と、母語が英語の現地の子どもの (n=1967) という3つのグループの得点を比較した。仮説では、文化負荷量と言語要求度が大きいとされる検査の得点が低く、その程度が小さいとされる検査の得点が高い傾向は、在米期間の短い子どものグループに一番強く表れ、次いで在米の文化的マイノリティ、そして、英語を第一言語とする子どもにはその傾向はないと考えられたが、グループ間

の傾斜パターンに明確な差がみられるという結果には至らなかった。

Styck と Watkins は 2 つの研究 (2013, 2014) を通じて、C-LIM が英語学習者と英語を第一言語とする子どもの得点とをどの程度正確に弁別できるかを Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV; Wechsler, 2003) の得点を用いて検証した。2013 年の研究では、6～16 歳の特別支援教育のニーズの有無を判定するためのアセスメントにリファーされた言語的マイノリティの英語学習者の子どもの得点 (n=86) を WISC-IV の標準化サンプルの得点と比較した。その結果、英語学習者の得点の方が、標準化サンプルの得点より有意に低く、得点のばらつきも大きかった。しかし、C-LIM が予測する得点の低下を示した英語学習者の子どもは 10.5%、標準化サンプルの子どもの方は 4.9%であった。

2014 年の研究では、6～16 歳の 3 つのグループの子どもの WISC-IV の得点を比較した。1 つ目のグループは特別支援教育のアセスメントにリファーされた言語マイノリティの子ども (n=69)、2 つ目は自閉症があるとされている現地の子ども (n=79)、3 つ目はアセスメントにリファーされたが、障害はないと評価された現地の子ども (n=216) である。1 つ目の研究 (Styck & Watkins 2013) と同様、グループの平均得点は、アセスメントにリファーされたが特別支援教育の対象ではないと判定された現地の子どもの得点よりも他の 2 つのグループ平均が有意に低くばらつきも大きいという結果となった。しかしながら、言語的マイノリティの子どもの 4.3%、自閉症の現地の子どもの 11.4%、特別支援教育の対象ではないと判定された英語を母語とする子どもの 24.4%しか C-LIM の予測する得点の低下パターンと一致しなかった。

これらの結果から、グループの平均的な得点の低下傾向のパターンは、あくまでも個々の得点が平均化された集団得点としての傾向であり、そのグループの個人個人にも、同様の傾向が当てはまるというわけではないということが考えられる (Styck & Watkins 2013; Calderon-tena et al., 2020)。このように C-LIM の妥当性については、その活用方法についてエビデンスが十分ではないとの指摘する研究が多い。これに対し、Ortiz &

Wong (2020) は、結果の解釈が C-LIM の意図と異なるため、これらの研究 (Styck & Watkins, 2013, 2014) で収集された得点傾向も C-LIM が示した得点傾向と一致していると反論している。C-LIM は、検査得点の妥当性を判断する際の参考基準であり、検査得点の低下が必ずしも障害の有無を示すものではないという理解で使われるべきものとしている (Flanagan et al., 2013)。

## 【考察とまとめ】

C-LIM を活用した得点の妥当性を確認する方法については、検査の標準実施を基本とするため、C-LIM の想定範囲外の得点低下がなければ、必ずしも子どもの母語版検査を実施できるバイリンガル検査者や通訳者が必要なく、著しい得点低下がみられた下位検査のみを再実施すればよいため、通訳者の負担が少ないという利点がある。ただし、例えば日本で行うことを想定した場合、日本語の教示が理解できない受検者には、事実上検査が成り立たず、C-LIM と照らし合わせる得点が算出できないというケースも想定しうる。Calderón-Tena et al. (2020) は、子どもの第二言語の習得が進んでいない場合には、非言語性検査を利用したり聞き取りなどを通して、質的情報を十分に生かしてアセスメントをすることが重要であると述べている。

Flanagan et al. (2013) が強調するように、C-LIM はあくまでも得点の妥当性を判断するための参考基準であるということは理解できる。しかしながら、C-LIM の示す得点傾向が個人得点の傾向に現れなければ、標準実施した下位検査得点において、文化負荷量と言語要求度が大きいとされる検査の得点が低く、その程度が小さいとされる検査の得点が高い傾向があるか否かを、文化と言語の違いによる影響の有無を見分けるための参考基準とすることは、現実には難しいのではないかとということが予想される。C-LIM の有用性に関する見解の不一致は、集団に見られる得点傾向と個人の得点傾向は必ずしも一致しないという研究結果 (e.g., Styck & Watkins, 2013; 2014) によって説明されるのが妥当であり、現段階では C-LIM の臨床的利用の具体的方法について、さらなる研究が必要であると考えられる。

一方、C-LIM の有用性について疑問を呈する研究 (e.g., Styck & Watkins, 2013) においても、

各検査の集団得点が概ね C-LIM の低下傾向と一致するという事は注目すべき点である。現時点では、C-LIM の臨床利用については限界があると思われるが、Naglieri & Otero (2017) で述べられているように、C-LIM は検査者が検査における文化的・言語的影響についての理解が不足している場合に、得点の傾向を参考にするという方法での活用の可能性はあると考えられる。

Hasson et al. (2019) の研究は文化グループによっては C-LIM の予測と一致しない低下度合いを示す下位検査があることを示唆している。また、米国版の C-LIM が予測する文化負荷量と言語要求度が日本版検査にも適合するという研究は現時点ではないため、本稿でまとめた C-LIM をすぐにわが国で活用することは考えにくい。しかしながら、文化と言語の影響を完全に切り分けることは難しいという視点はあるものの (Cormier et al., 2014)、検査における文化負荷量と言語要求度という 1 つの考え方の枠組みを示しているという点において、文化や言語の背景が異なる子どもの検査得点を理解する際の一助となると考えられる。また、文化と言語の影響によって検査得点が低くなる傾向は、言語性か非言語性か等の下位指標レベルを超えて、下位検査ごとに異なる可能性があるということを示している点についても、わが国の外国にルーツをもつ子どものアセスメント研究を進める場合に参考になると考えられる。

Cormier et al. (2014) と Hasson et al. (2019) の結果は、文化負荷量に関して適切な指標がなく、その推定は容易ではないことを示唆している。最近の研究 (e.g., Cormier et al., 2022) では、これまでの研究の多くが検査の特徴に注目してきたが、研究が進み受検者側の特徴、特に言語力を加えたモデルの検証が可能になっていると述べている。社会経済的背景 (Socio-Economic Status: SES) が知能検査の得点に与える影響についての研究もある (e.g., Hasson et al. 2019; Hanscombe, 2012)。日本における外国にルーツをもつ子どもは文化や言語の違いだけではなく、移動の背景にある経済的状況や貧困の問題などの影響をうけている場合がある (金, 2020)。このような点も踏まえて、国外での研究の動向にも注目しつつ、わが国でのアセスメントの方法について検討をしていく必要があるであろう。

## 【文献】

- Calderón-Tena, C. O., Styck, K. Vega, D., Kranzler, J. H. (2020). Evaluating cultural and linguistic load of IQ scores for English language learners. *International Journal of School & Educational Psychology*, 46-62.
- Cormier, D. C., Bukut, O., McGrew, K. S., Kennedy, K. E. (2022). Linguistic Influences on Cognitive Test Performance: Examinee Characteristics Are More Important than Test Characteristics. *Journal of Intelligence*, 10, 8.
- Cormier, D. C., McGrew, K. S., & Ysseldyke, J. E. (2014). The Influences of Linguistic Demand and Cultural Loading on Cognitive Test Scores. *Journal of Psychoeducational Assessment*, 32, 610-623.
- Cormier, D. C., Wang, K., & Kennedy, K. E. (2016). Linguistic Demands of the Oral Directions for Administering the WISC-IV and WISC-V. *Canadian Journal of School Psychology*, 31, 290-304.
- Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. College-Hill Press, San Diego.
- Flanagan, D. P. & Alfonso, V. C. (2011). *Essentials of specific learning disability identification*. 上野一彦, 名越斉子監訳 (2013) エッセンシャルズ—新しい LD の判断. 日本文化科学社.
- Flanagan, D. P. & Alfonso, V. C. (2017). *Essentials of WISC-V assessment*. Wiley, Hoboken.
- Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2007). *Essentials of cross-battery assessment* (2nd ed.). Wiley, Hoboken.
- Flanagan, D. P., Ortiz, S. O., & Alfonso, V. C. (2013). *Essentials of cross-battery assessment* (3rd ed.). Wiley, Hoboken.
- Hanscombe, B. K., Trzaskowski, M., Haworth, C. M. A., Davis, O. P. S., Dale, P. S., & Plomin, R. (2012). Socioeconomic status (SES) and children's intelligence (IQ): In a UK-representative sample SES moderates the environmental, not genetic, effect on IQ. *PLOS ONE*, 7, 1-16.
- Hasson R., Wu L., & Fine J. (2019). Clinical utility of the WASI-II and its association with acculturation levels among Arab American adolescent males. *Applied Neuropsychology: Child*. 2019, 8, 295-306.
- Herrera, S.G., Cabral, R. M., & Murry, K. G. (2020). *Assessment of Culturally and Linguistically Diverse Students*. Pearson, Hoboken.
- 法務省入国管理局 (2021). 令和 3 年末現在における在留外国人数について (確定値), 法務省. <https://www.immi.go.jp/>

- //www.moj.go.jp/isa/content/001370057.pdf (2022年1月6日閲覧)
- Jacob, S. & Hartshorne, T. S. (2003). *Ethics and law for school Psychologists* (4th ed.). Wiley, Hoboken.
- Kaufman, A. S. & Kaufman, N. L. (2004a). *Kaufman assessment battery for children, second edition*. AGS Publishing, Circle Pines.
- Kaufman, A. S. & Kaufman, N. L. (2004b). *Kaufman test of educational achievement, second edition*. Pearson, San Antonio.
- Kaufman, A. S. & Kaufman, N. L. (2018). *Kaufman assessment battery for children, second edition, normative update*. NCS Pearson, Bloomington.
- Kaufman, A. S., Lichtenberger, E. O., Fletcher-Janzen, E., & Kaufman, N. L. (2005). *Essentials of KABC-II Assessment*. Wiley, Hoboken.
- 金春喜 (2020) 「発達障害」とされる外国人の子どもたちフィリピンから来日したきょうだいをめぐり、10人の大人たちの語り。赤石書店。
- Kranzler, J. H., Flores, C. G., & Coady, M. (2010). Examination of the cross-battery approach for the cognitive assessment of children and youth from diverse linguistic and cultural backgrounds. *School Psychology Review*, 39, 431-446.
- 前川久男, 中山健, 岡崎慎治 (2007a). 日本版 DN-CAS 認知評価システム。日本文化科学社。
- 前川久男, 中山健, 岡崎慎治 (2007b). 日本版 DN-CAS 理論と解釈のためのハンドブック。日本文化科学社。
- 松田真希子・中川郷子 (2017). 外国にルーツをもつ児童の発達アセスメントと言語の問題について—発達障害と一時的リミテッド状況の鑑別のための調査研究, 金沢大学留学生センター紀要, 21, 29-42.
- 丸山宏樹 (2013). 学習や行動上の困難を呈する外国人の子どもの特徴—WISC-IIIの結果の分析からの検討—日本教育心理学会総会発表論文集 55, 325.
- 三好一英・服部環 (2010). 海外における知能研究とCHC理論. *Tsukuba Psychological Research*, 40.
- 文部科学省総合教育政策局 (2019) 外国人児童生徒受け入れの手引き改訂版 [https://www.mext.go.jp/a\\_menu/shotou/clarinet/002/1304668.htm](https://www.mext.go.jp/a_menu/shotou/clarinet/002/1304668.htm) (2023年1月6日閲覧)
- Naglieri, J. A. & Das, J. P. (1997). *Cognitive assessment system*. Riverside, Itasca.
- Naglieri, J. A., Das, J. P., & Goldstein, S. (2014). *Cognitive assessment system, second edition*. Pro-Ed, Austin.
- Naglieri, J. A. & Otero, T. M. (2017). *Essentials of CAS2 assessment*. Wiley, Hoboken.
- 二井紀美子・緩利 誠 (2013). 外国人児童生徒支援に資するアセスメントの枠組みの提案—不就学児調査を通して—生涯学習・キャリア教育研究, 9, 1-12.
- 日本版 KABC-II 制作委員会訳編 (2013a). 日本版 Kaufman assessment battery for children second edition: KABC-II. 丸善出版.
- 日本版 KABC-II 制作委員会訳編 (2013b). 日本版 KABC-II 制作委員会訳編 日本版 KABC-II マニュアル. 丸善出版.
- 日本版 WISC-IV 刊行委員会訳編 (2010a). 日本版 WISC-IV 知能検査. 日本文化科学社.
- 日本版 WISC-IV 刊行委員会訳編 (2010b). 日本版 WISC-IV 知能検査理論解釈マニュアル. 日本文化科学社.
- 日本版 WISC-V 刊行委員会訳編 (2022a). 日本版 WISC-IV 知能検査. 日本文化科学社.
- 日本版 WISC-V 刊行委員会訳編 (2022b). 日本版 WISC-IV 知能検査実施・採点マニュアル. 日本文化科学社.
- Ortiz, S. (2017). Multicultural issues in school neuropsychological evaluation: Best practices in nondiscriminatory assessment of culturally and linguistically diverse populations. 2016-2017 school neuropsychology post-graduate training program.
- Ortiz, S. O. (2023). Culture-language interpretive matrix-Basic version v5.0 (Excel) (updated 2.1.23) <http://facpub.stjohns.edu/~ortiz/CLIM/> (2023年2月19日閲覧)
- Ortiz, S. O. & Wong, J. Y. T. (2020) Bilingual evaluation for non-bilingual school psychologists: an evidence-based approach to evaluation and best practice. National Association of school psychologists' annual convention. Baltimore, MD.
- Rhodes, R. L., Ochoa, S. H., & Ortiz, S. O. (2005). *Assessing culturally and linguistically diverse students: A practical guide*. Guilford Press, New York.
- Ryan, J. J., Glass, L.A., & Bartels, J. M. (2010). Stability of the WISC-IV in a sample of elementary and middle school children, *Applied Neuropsychology*, 17, 68-72.
- Roid, G. H. (2003). *Stanford-Binet intelligence test-fifth edition*. Riverside, Itasca.
- 島田直子 (2016). 多文化背景の子どもたちへの知能検査の利用法—心理教育アセスメントに関する近年の米国の文献から—LD研究, 25, 358-367.
- 荘島宏二郎 (2014). 学力を図る. 日本発達心理学会 (編) 発達心理学辞典, pp. 522-523.
- Sotelo-Dynega, M. & Dixon, S. G. (2014). Cognitive assessment practices: A survey of school psycholo-

- gists. *Psychology in the Schools*, 51, 1031-1045.
- Styck, K. M. & Watkins, M. W. (2013). Diagnostic utility of the culture and language interpretation matrix for the WISC-IV among referred students. *School Psychology Review*, 42, 367-382.
- Styck, K. M. & Watkins, M. W. (2014). Discriminant validity of the WISC-IV culture and language interpretation matrix. *Contemporary School Psychology*, 18, 168-177.
- 杉原一昭, 杉原隆監修, 中村淳子, 大川一郎, 野原理恵ら編 (2005). 田中ビネー知能検査V理論マニュアル. 田研出版.
- 田中教育研究所 (2005). 田中ビネー知能検査V. 田研出版.
- Wechsler, D. (2004). *Wechsler intelligence scale for children-fourth edition*. Pearson, San Antonio.
- Wechsler, D. (2014). *Wechsler intelligence scale for children-fifth edition*. Pearson, San Antonio.
- Wechsler, D. (2017). *Wechsler intelligence scale for children-fifth edition Spanish*. Pearson, San Antonio.
- 緩利 誠・二井紀美子 (2010). 不就学外国人児童にみられる知的能力の実態とその傾向－WISC-IIIの結果分析を中心に－日本教育学会大会研究発表要項 69, 390-391.