

【原著】

WISC- IV 下位検査内のばらつき (ISS) に 関する基準値の作成

佐藤 匠 神田 聡 舘花佳奈子

〈要旨〉本研究はWISC- IVの各基本下位検査内における正答と誤答のばらつき（以下ISS）の基準値を検証した。370名の児童を対象とし、341名はASD、AD/HD、SLD、ID、DCDいずれかの診断を有し、29名は診断がつかなかった。各下位検査の中止条件の対象となる誤答を除いた誤答数の平均値、標準偏差を算出した。誤答数の平均値よりも1標準偏差大きい場合にはISSが大きいものと見なした。その結果、類似は4個、積木模様は2個、絵の概念は6個、行列推理は4個、順唱と逆唱は2個、符号は2個が基準値と考えられた。単語、理解、語音整列は年齢によってISSの基準値が変化し、記号探しはFSIQ水準によって基準値が変化すると考えられた。

キーワード：WISC- IV、下位検査内のばらつき、アセスメント

1. 問題と背景

WISC- IVでは評価点を加味した合成得点中心の解釈が推奨されており、下位検査内の評価点レベルの量的、質的情報については単独で解釈をしていくのではなく、複数の情報と統合した解釈が必要とされている（松田,2015）。評価点レベルの量的、質的情報を他の情報と統合していく手法の1つとして、得点に至る問題解決方略や誤答の背景にある過程を理解するプロセスアプローチ（Kaplan,1988）が挙げられる。海外で使用されているWISC- IV統合版では多くのプロセスアプローチが適用されており、児童がどのような認知方略を使用しているのかを検討するために、通常のWISC- IVとは異なった下位検査を用意している。例えば言語理解の下位検査では回答を選択肢から選ぶ多肢選択式版が用意されており、検査場面の「分からない」という回答や問題を繰り返し

聞き返すこと等の反応の累積率を算出できるようになっている（McCloskey & Maerlender,2005）。McCloskey et al.(2005) はWISC- IV統合版における追加的なプロセスアプローチ下位検査を実施することによって、認知的な強みと苦手さに対する理解が深まった事例を紹介している。WISC- IV統合版の各種指標得点等は児童の認知能力や課題方略の理解を助け、仮説生成や他の検査実施の判断を支えるものと考えられる。プロセスアプローチと類似した手法のCHT(Cognitive Hypothesis Testing)modelでは、検査で要求される能力を入力—処理—出力と情報処理の観点から分類し、児童の強みと苦手さに関する仮説を立てて検証していく（Hale & Fiorello,2004）。Miller & Jones(2016) は、WISC- IVの改訂版にあたるWISC- Vに対して、神経心理学とCHC理論を統合したSNP/CHC modelから解釈する方法を提案している。SNP/CHC modelの手法では

Takumi Sato：島田療育センターはちおうじ
Satoshi Kanda：島田療育センターはちおうじ
Kanao Tatchana：島田療育センターはちおうじ

WISC- VやWISC- V統合版の下位検査に対して4領域(基本的な感覚運動機能,認知処理と獲得した知識スキルの促進と抑制,基本的な認知処理,獲得した知識)の観点から分類を行い、発達障害が認知処理に及ぼす影響の評価などを目的に用いられている(Miller et al.,2016)。いずれの手法も単一の検査ではなく、国内では翻訳や開発されていない多くの検査を用いて検証していくことが求められており、同様の手続きで取り組むことは困難である。しかしこれらの手法は児童の強みと苦手さを把握する上で有益なものであり、解釈の参考となる部分が多い。国内においても児童の検査場面の取り組み等、質的情報を判断し、複数の情報と統合させていきながら解釈していく方法を確立していくことが必要と考えられる。

検査の取り組み方に関する質的情報の判断として、岡田・田邊・飯利(2015)は検査場面における行動観察のアセスメントとして「意欲・協力的態度」「知的理解・言語理解」「不注意・衝動性」「社会性の困難」「こだわり・切り替えの困難」「情緒の問題」と6領域から構成される検査行動チェックリストを開発している。検査場面で見られる特性を幅広く捉えることが可能な検査行動チェックリストであり、該当する項目数が多いようであれば追加の検査や日常生活の様子を聴取する一つの判断基準となる。竹厚・池田・山内(2014)は絵の抹消の抹消手順の分類をし、知的水準毎にどのような抹消手順が多く見られたかを示しており、1つの質的分析の判断基準となる。

その他の検査場面の質的情報として下位検査内の正答と誤答のばらつきが挙げられ、Intra-subtest Scatter(以下ISS)と呼ばれている(Godber,Anderson & Bell,2000; Courville, Coalson & Kaufman et al.,2016)。ISSを算出する幾つかの方法が開発されており、Juni & Troblier(2009)は近年の標準的な方法として、連続する2つの項目得点間の絶対差の総和を用いる方法を挙げている。これは下位検査内の得点が1の後に0、もしくは0の後に1と変化した場合に「1」と得点する、2の後に0、もしくは0の後に2と変化した場合に「2」と得点する方法である(Dumont

& Willis,1995)。Dumont et al.(1995)は健常群と音声言語障害やADHD等の障害群において、WISC- IIIの下位検査毎のISSを比較した結果、両群の違いが見られなかったことを報告している。Courville et al.(2016)は同様の方法をWISC- Vでも検証したところ、健常群と臨床群に違いが見られなかったことを報告している。Juni et al.(2009)も健常群と障害群の違いが見られなかったことを報告しており、ISSが障害群と臨床群の違いを示すものではないと考えられる。ただ、これらの研究は健常と障害の判別に焦点が当てられており、ISS自体は児童の課題パフォーマンス向上のための知見を得る上で有効という指摘がある(McCloskey et al.,2005)。ISSは障害の判別を目的に解釈することやISSの情報単独で解釈するのではなく、解釈の仮説を立て、他の情報と合わせて検証していくように活用することで有益な情報となる可能性が考えられる。特に検査前の事前情報が不足していたり、検査場面では刺激が少なく、手続きが構造化されているために岡田ら(2015)の検査行動チェックリストで特徴が見られなかったりする場合に、ISSから解釈の仮説を立てることが理解を深める手がかりとなる事例が存在する可能性が考えられる。

ISSを算出する、連続する2つの項目得点間の絶対差の総和を用いる方法は、どこの問題で間違えたかという情報が反映されない課題を踏まえ、項目-応答理論を用いたGodber et al. (2000)の手法や下位検査内の得点のばらつきを11の水準に分類したJuni et al.(2009)の手法が開発されている。これらは言語理解や知覚統合、知覚推理の下位検査を対象としているため、いずれの手法もワーキングメモリー、処理速度の下位検査を対象としていない。またJuni et al.(2009)の方法は検査中の判断が難しいものと考えられ、ISSは単独で解釈していくのではなく、その他の情報と統合して考えていく性質のものと考え、検査後の聴取や補助検査の必要性をすぐに判断できるものが望ましい。検査者がISSの判断が容易に行える手法を開発することで、検査場面で補助検査の実

施判断や、検査後に保護者へ日常生活のことを聴取する判断の助けとなる可能性が考えられる。検査者は通常中止条件を判断するために誤答数を把握しながら取り組んでいることが求められるため、誤答数の把握の方が容易と考えられる。誤答数の平均値、中央値や標準偏差が分かることでISSの大きさを判断することが可能となり、ISSにどの程度の重きを置いて仮説検討するかの判断材料となると考えられる。障害群と健常群でISSの違いはないため (Courville et al.,2016;Juni et al.,2009;Dumont et al.,1995)、特定の点数を超えた場合に認知特性や障害特性を反映していると考える基準値ではなく、統計的に少ない発生率をISSが大きい基準値として適切と考えられる。本研究では誤答数から1標準偏差を超える場合にはISSが大きいものと見なし、1標準偏差を下回る場合にはISSが小さいものと見なす。またISSが大きい人数、ISSが小さい人数を算出し、人数比の比較を行う。その結果から、1標準偏差を基準とするよりも、より大きいもしくは少ない基準値でISSの大きさを検討するか検証する。

そこで本研究はWISC- IVの基本検査内の誤答数を対象とし、平均値、標準偏差、中央値、年齢やIQによる違いを検証し、WISC- IVの下位検査内のISSの基準値の作成を行う。

なお本研究では「数唱」のプロセス得点である「順唱」と「逆唱」はそれぞれ評価点があり、単独で評価することが臨床的に価値があると考え、「順唱」「逆唱」を「数唱」の代わりの下位検査と見なす。

2. 方法

2.1 対象

20XX年11月～20XX+4年3月までA医療機関を受診した370名の児童(平均年齢9歳6ヶ月 \pm 2.8SD、年齢範囲:5歳0ヶ月～16歳11ヶ月)を対象とした。対象児童は医師によるDSM-5に基づいたASD(自閉症スペクトラム障害)、AD/HD(注意欠如・多動性障害)、SLD(限局性学習障

害)、ID知的発達障害)、DCD(発達性協調運動障害) いずれかの診断を受けているものが大半であった。児童の診断の内訳として、ASD=78名、AD/HD=79名、SLD=18名、ID=19名、ASD+AD/HD=46名、ASD+ID=26名、ASD+DCD=11名、ASD+SLD=7名、AD/HD+SLD=13名、AD/HD+DCD=14名、ASD+AD/HD+DCD=8名であった。DCD、AD/HD+ID、DCD+ID、SLD+DCD、ASD+AD/HD+SLD、ASD+AD/HD+ID、ASD+SLD+ID、AD/HD+DCD+ID、AD/HD+SLD+ID、AD/HD+SLD+DCDの診断がついているものはいずれも5名以下であり、これらの診断がついた児童の総数は22名であった。診断がついていない児童は29名であった。児童全体の検査結果はFSIQ=87.54(14.05)、VCI=89.76(14.80)、PRI=89.58(15.00)、WMI=87.63(16.04)、PSI=91.57(14.48)であった。下位検査で粗点が0点だった児童は19人(「類似」9人、「理解」1人、「逆唱」5人、「語音整列」4人)いたが、同一人物内で粗点が0点となる複数の検査が存在することはなかったため研究の対象とした。

2.2 倫理手続き

本研究を実施するに当たり、社会福祉法人島田療育センターはちおうじ倫理審査委員会の承認を得た上で、全ての保護者に口頭での説明並びに研究承諾書の書面にて同意のサインを得られた児童のみを対象とした。

2.3 誤答数の数え方

本研究の下位検査内の誤答数の数え方を表1に示す。

表1 誤答の数え方

【誤答の数え方のルール】

- 1: 誤答(0点)の数を数え、合計数を算出する
- 2: 中止条件に含まれる誤答は誤答数として数えない
- 3: 最後の問まで到達し、中止条件に達していない場合は、最後の正答が得られた問題に至るまでの誤答を対象とする。
- 4: 練習問題は誤答として見なさない。

【行列推理を例として】

1: 問11～22まで1点が連続で得られ、問23で0点、問24で1点、問25～27で連続0点の場合

→問23の誤答(0点)は中止条件の対象となるため、誤答と見なさない。よって、この場合の誤答数は0個である。

2: 問11～22まで1点が連続で得られ、問23～24で連続0点、問25～26で連続1点、問27～30で連続0点の場合

→問23の0点、問24の0点が誤答となり、誤答数は2個である。

3: 問11～30まで連続で1点、問31～32で連続0点、問33で1点、最後の問34で0点

→最後の問まで到達し、中止条件に達していない。最後に得られた正答が問33のため、問31～32の誤答が対象となり、誤答数は2個である。

中止条件の対象となる誤答以外の誤答を数え、総数を算出した。最後の問題まで到達し、中止条件に至らなかった場合は、最後の正答が得られた問題までの誤答を対象として見なした。練習問題の誤答は数として含めなかった。例えば行列推理で問11～22まで正答が得られ、問23で誤答、問24で正答、問25～27で誤答となった場合には誤答数を0と数えた。問11～22まで正答が得られ、問23～24で誤答、問25～26で正答、問27～30で誤答となった場合は誤答数を2と数えた。問11～30まで正答、問31～32で誤答、問33で正答、最後の問34で誤答となった場合は誤答数2と数えた。符号は書き間違えのミスのみを誤答の対象とし、記号探しは間違いのみを対象とし、問題を飛ばすことや自発的修正は誤答の対象としなかった。

2.4 分析方法

まずWISC-IVの基本検査を対象に誤答数を算出し、研究対象児童の平均値、標準偏差、中央値、最小値、最大値を算出した。ただし下位検査の粗点が0点だったものは該当する下位検査の分析を行う場合のみ除外した。例えば「類似」で粗点が0点だった場合に「類似」の誤答数を検討する場合には分析から除外するが、他の下位検査の分析では使用した。

本研究ではISSの大きい児童の基準を1標準偏差とした。ISSの小さい基準は「誤答数の平均値－標準偏差」とし、ISSの大きい基準は「誤答数の平均値＋標準偏差」とした。平均値から標準偏差を引いた場合に小数点が見られた場合は1繰り下げること、平均値から標準偏差を足した合計値で小数点が見られた場合には1繰り上げることとした。この基準値が年齢毎に違うかを検討するため、5歳～16歳において1歳毎に平均値、標準偏差を算出した。ただし14歳～16歳の人数は少なかったため、14歳～16歳を1つの群として見なした。その後、年齢間の誤答数の比較をするために一要因分散分析を行った。続いてFSIQ水準毎の違いを検討するため「FSIQ≤70」「FSIQ:71~85」「FSIQ:86~100」「FSIQ:101~115」「FSIQ>115」と1標準偏差毎に群を作成した。FSIQ群毎に平均値、標準偏差を算出した。続いてFSIQ群毎の誤答数を比較するために一要因分散分析を行った。いずれも有意差が見られた場合には多重比較(Shaffer法)を行った。効果量の算出はCohen's $d = (\bar{X}_1 - \bar{X}_2) / \sqrt{((n_1 - 1)S_1^2 + (n_2 - 1)S_2^2) / (n_1 + n_2 - 2)}$ を用いた(大久保, 岡田, 2012)ものの、Elis(2010)は効果量の解釈に関しては該当する研究分野によって解釈の基準が異なると指摘している。本研究では効果量は算出するものの、実際のISSの判断基準値、すなわち誤答数の判断基準値が変わるかの判断に重きを置いた。年齢間、FSIQ間で有意差が複数の群間で見られた場合には、有意差が見られた群や各年齢、FSIQ群における誤答数の基準値などを参考に群を作成し、誤答数の基準値を検証した。

3. 結果

3.1 各下位検査における誤答数の基本統計量

表2に各下位検査における誤答数の平均値、標準偏差、中央値、最小値、最大値を示す。なお粗点が0点の場合には、該当する下位検査から除外して算出した。

表2 下位検査における誤答数の平均値、標準偏差、中央値、最大値、最小値

	類似	単語	理解	積木模様	絵の概念	行列推理	順唱	逆唱	語音整列	符号	記号探し
Mean	1.99	2.81	2.05	0.71	3.22	1.62	0.76	0.91	3.01	0.48	1.54
SD	1.63	2.51	1.58	0.87	2.32	1.58	0.75	0.87	2.14	1.03	1.80
Mdn	2	2	2	0	3	1	1	1	3	0	1
最小値	0	0	0	0	0	0	0	0	0	0	0
最大値	9	12	7	4	13	8	3	4	11	9	9
総人数	361	370	369	370	370	370	370	365	366	370	370

*粗点が0点の被験者は除いたため、類似、理解、逆唱、語音整列の被験者は370人よりも少ない人数となっている。

各下位検査の誤答数の平均値(標準偏差)は以下の通りであった。「類似」=1.99(1.63)、「単語」=2.81(2.51)、「理解」=2.05(1.58)、「積木模様」=0.71(0.87)、「絵の概念」=3.22(2.32)、「行列推理」=1.62(1.58)、「順唱」=0.76(0.75)、「逆唱」=0.91(0.87)、「語音整列」=3.01(2.14)、「符号」=0.48(1.03)、「記号探し」=1.54(1.80)であった。各下位検査の平均値から1標準偏差を足した場合はISSが大きいと判断し、引いた場合にはISSが小さいと判断した。また平均値から標準偏差を足した時に小数点が見られた場合には1繰り上げた誤答数を基準とし、平均値から標準偏差を引いた場合に小数点が見られた場合には1繰り下げた誤答数を基準とした。例えば「類似」の平均値=1.99、標準偏差=1.63である。平均値に標準偏差を足すと3.62となり、小数が見られるため1繰り上げる処理を行うと「4」になり、誤答数の多い基準値は4となる。

これらの処理を行ったところ「類似」は4個、「単語」は6個、「理解」は4個、「積木模様」は2個、「絵の概念」は6個、「行列推理」は4個、「順唱」は2個、「逆唱」は2個、「符号」は2個、「記号探し」は4個、これらの数値以上の場合にはISSが大きいものと見なした。ISSが小さい基準はいずれの下位検査も0個の場合と考えられた。誤答数が0個の場合には下位検査の誤答の偏りが無く、ISSが小さいと表現するよりもISSの無い群と表現する方が適切と考えられる。そのため、これ以降はISSの小さい群ではなく、ISSの無い群の人数を算出する。

3.2 年齢毎の各下位検査の誤答数の比較

各年齢の誤答数の平均値、標準偏差を表3に示す。なお「類似」は粗点が0点の被験者が9人いたため、5歳群は24人、6歳群42人、8歳群50人であった。「理解」は粗点が0点の被験者が1人いたため、6歳群46人であった。「逆唱」は粗点が0点の被験者が5人いたため、5歳群は23人、6歳群は46人、7歳群は57人、8歳群は52人であった。「語音整列」は粗点が0点の被験者が4人いたため、5歳群は24人、6歳群は46人、7歳群は57人、8歳群は52人であった。

年齢による誤答数の違いを下位検査毎に検証した一要因分散分析の結果を表4に示す。

その結果「類似」: $F(9,351)=2.20, p=.02, \eta^2=.05$ 「積木模様」: $F(9,360)=1.20, p=.30, \eta^2=.03$ 「絵の概念」: $F(9,360)=1.44, p=.17, \eta^2=.03$ 「行列推理」: $F(9,360)=2.10, p=.03, \eta^2=.05$ 「順唱」: $F(9,360)=0.54, p=.85, \eta^2=.01$ 「逆唱」: $F(9,355)=1.85, p=.06, \eta^2=.04$ 「符号」: $F(9,360)=0.51, p=.86, \eta^2=.01$ 「記号探し」: $F(9,360)=0.16, p=.99, \eta^2<.01$ では有意差が見られなかった。

「単語」: $F(9,360)=10.91, p<.01, \eta^2=.21$ 「理解」: $F(9,359)=2.06, p=.03, \eta^2=.04$ 「語音整列」: $F(9,356)=7.56, p<.01, \eta^2=.16$ では5%水準で有意差が見られたため、多重比較(Shaffer法)を行い、5%水準で有意差が見られた年齢間のみを表4に示す。「単語」では10歳、11歳、12歳、13歳、14歳以上の群が5歳、6歳、7歳の群よりも誤答数が多い結果となり、効果量は $d=0.89\sim 1.88$ の範囲に位置した。10歳、11歳、13歳、14歳以上の群は8歳群よりも誤答数が多く、効果量は $d=0.76$

表3 各年齢における下位検査の誤答数数の平均値、標準偏差

	類似	単語	理解	積木模様	絵の概念	行列推理	順唱	逆唱	語音整列	符号	記号探し
5歳	1.45	1.52	1.16	0.80	2.16	0.84	0.80	0.65	1.00	0.32	1.60
25人	(1.14)	(1.36)	(1.28)	(0.72)	(2.08)	(1.46)	(0.82)	(0.78)	(1.18)	(0.75)	(1.50)
6歳	1.71	1.43	1.97	0.70	3.43	1.17	0.74	0.70	1.93	0.66	1.64
47人	(1.38)	(1.68)	(1.42)	(0.75)	(2.52)	(1.22)	(0.74)	(0.76)	(1.81)	(1.56)	(1.85)
7歳	1.55	1.88	1.78	0.81	3.31	1.47	0.66	0.74	2.69	0.60	1.52
58人	(1.31)	(1.62)	(1.50)	(0.91)	(2.15)	(1.57)	(0.64)	(0.64)	(2.14)	(1.32)	(2.21)
8歳	2.04	1.98	1.91	0.79	3.51	1.60	0.70	1.06	3.62	0.45	1.47
53人	(1.47)	(2.17)	(1.54)	(0.95)	(2.02)	(1.46)	(0.72)	(0.85)	(2.17)	(0.85)	(1.56)
9歳	2.47	2.79	2.00	0.81	2.63	2.02	0.72	0.88	2.58	0.56	1.51
43人	(2.14)	(2.48)	(1.73)	(1.05)	(1.70)	(1.85)	(0.73)	(0.91)	(1.74)	(0.91)	(2.21)
10歳	2.00	4.00	2.29	0.74	3.61	1.58	0.90	1.19	3.35	0.42	1.35
31人	(1.51)	(3.02)	(1.66)	(0.93)	(3.31)	(1.52)	(0.87)	(0.91)	(1.92)	(0.85)	(1.58)
11歳	2.36	3.86	2.50	0.48	3.59	1.89	0.77	0.93	3.91	0.39	1.48
44人	(1.4)	(2.82)	(1.53)	(0.70)	(2.60)	(1.62)	(0.83)	(1.04)	(2.39)	(0.75)	(1.96)
12歳	2.48	4.04	2.28	0.76	3.44	2.12	0.96	1.08	3.92	0.32	1.88
25人	(2.49)	(2.17)	(1.70)	(1.13)	(2.60)	(1.86)	(0.79)	(0.86)	(1.96)	(0.69)	(1.69)
13歳	2.44	4.89	2.33	0.50	3.22	1.78	0.83	1.28	4.00	0.50	1.50
18人	(1.15)	(2.21)	(1.82)	(0.62)	(2.16)	(1.70)	(0.86)	(1.18)	(1.91)	(0.86)	(1.82)
14歳以上	1.58	4.50	2.54	0.38	2.70	1.96	0.73	0.85	3.54	0.35	1.54
26人	(1.80)	(2.79)	(1.58)	(0.64)	(1.72)	(1.43)	(0.60)	(0.78)	(1.90)	(0.85)	(1.39)

*上段が平均値、下段が標準偏差である。

*類似は粗点が0点の被験者が9人いたため、5歳は24人、6歳42人、8歳50人であった。

*理解は粗点が0点の被験者が1人いたため、6歳46人であった。

*逆唱は粗点が0点の被験者が5人いたため、5歳は23人、6歳は46人、7歳は57人、8歳は52人であった。

*語音整列は粗点が0点の被験者が4人いたため、5歳は24人、6歳は46人、7歳は57人、8歳は52人であった。

表4 各年齢間の下位検査の誤答数を比較した分散分析の結果

下位検査	F値(df)	多重比較(5%水準で有意差が見られたもの)
類似	F(9,351)=2.20,p=.02,η ² =.05	
単語	F(9,360)=10.91,p<.01,η ² =.21	14歳以上,13歳,11歳,10歳,12歳>6歳(d=1.06~1.88) 13歳,14歳以上,11歳,10歳,12歳>7歳(d=0.89~1.70) 13歳,11歳,10歳,12歳,14歳以上>5歳(d=0.97~1.91) 13歳,14歳以上,11歳,10歳>8歳(d=0.76~1.44) 13歳>9歳(d=0.87)
理解	F(9,359)=2.06,p=.03,η ² =.49	11歳>5歳(d=0.93)
積木模様	F(9,360)=1.20,p=.30,η ² =.03	
絵の概念	F(9,360)=1.44,p=.17,η ² =.03	
行列推理	F(9,360)=2.10,p=.03,η ² =.05	
順唱	F(9,360)=0.54,p=.85,η ² =.01	
逆唱	F(9,355)=1.85,p=.06,η ² =.04	
語音整列	F(9,356)=7.56,p<.01,η ² =.16	11歳,8歳,12歳,13歳,14歳以上,10歳,7歳>5歳(d=0.89~1.97) 11歳,8歳,12歳,13歳,14歳以上>6歳(d=0.84~1.13)
符号	F(9,360)=0.51,p=.86,η ² =.01	
記号探し	F(9,360)=0.16,p=.99,η ² <.01	

～1.44の範囲に位置した。13歳群に関しては9歳群よりも誤答数が多く、効果量はd=0.87であった。「単語」における各年齢群の平均値と標準偏差を足して誤答数の判断基準値を算出したとこ

ろ、5歳群：3個、6歳群～7歳群：4個、8歳群：5個、9歳群：6個、10歳群：8個、11歳群：7個、12歳群：7個、13歳群～14歳以上群：8個であった。5歳群～9歳群と10歳群～14歳以上の群の間で

ISSの平均値の差が多く見られたと考え、5歳から9歳までを1つの群、10歳以上を1つの群としてISSの平均値、標準偏差を算出した。その結果5歳から9歳の平均値=1.94、標準偏差=1.98、平均値に1標準偏差を加えた値は3.92のため、5歳から9歳はISSの基準値は4と見なした。10歳以上の平均値=4.17、標準偏差=2.67、平均値に1標準偏差を加えた値は6.84のため、10歳以上はISSの基準値は7と見なした。

「理解」は11歳群が5歳群よりも誤答数が多く、効果量は $d=0.93$ であった。「理解」における各年齢群の平均値と標準偏差を足して誤答数の判断基準値を算出したところ、5歳群:3個、6歳群~10歳群:4個、11歳群:5個、12歳群:4個、13歳群:5個、14歳以上群:5個であった。「理解」の6歳群から14歳以上の群のISS基準値は4個から5個のため、5歳を除いた「理解」のISSの平均値と標準偏差を算出したところ、平均値=2.11、標準偏差=1.56なのであった。平均値に1標準偏差を加えた値は3.67のため、6歳以上はISSの基準値は4と見なした。

「語音整列」では7歳、8歳、10歳、11歳、12歳、13歳、14歳以上の群は5歳群よりも誤答数が多く、効果量は $d=0.89 \sim 1.97$ の範囲に位置した。8歳、11歳、12歳、13歳、14歳以上の群は6歳群よりも誤答数が多く、効果量は $d=0.84 \sim 1.13$ の範囲

に位置した。「語音整列」における各年齢群の平均値と標準偏差を足して誤答数の判断基準値を算出したところ、5歳群:3個、6歳群:4個、7歳群:5個、8歳群:6個、9歳群:5個、10歳群:6個、11歳群:7個、12歳群:6個、13歳群:6個、14歳以上群:6個であった。「語音整列」は5歳、6歳が多くの年齢群よりもISSの平均値が低い結果が出ているため、5歳と6歳を1つの群、7歳以上を1つの群に分け、ISSの平均値と標準偏差を算出した。その結果5歳から6歳群の平均値=1.57、標準偏差=1.68であり、平均値に1標準偏差を加えた値は3.25のため、5歳から6歳はISSの基準値は4と見なした。7歳以上の平均値=3.32、標準偏差=2.21であり、平均値に1標準偏差を加えた値は5.53のため、7歳以上はISSの基準値は6と見なした。

3.3 FSIQ 水準毎の各下位検査の誤答数の比較

各FSIQ水準の誤答数の平均値、標準偏差を表5に示す。

「類似」は粗点が0点の被験者が9人いたため、FSIQ: ≤ 70 群は39人、FSIQ:71~85群は132人、FSIQ:86~100群は127人であった。「理解」は粗点が0点の被験者が1人いたため、FSIQ:86~100群は129人であった。「逆唱」は粗点が0

表5 各IQ水準における下位検査の誤答数の平均値、標準偏差

FSIQ	類似	単語	理解	積木模様	絵の概念	行列推理	順唱	逆唱	語音整列	符号	記号探し
≤ 70	1.95	2.20	2.15	0.60	3.38	1.43	0.68	0.90	2.28	0.65	1.85
40人	(1.45)	(2.16)	(1.58)	(0.84)	(2.44)	(1.58)	(0.66)	(0.85)	(2.30)	(1.41)	(1.82)
71~85	2.11	3.03	2.07	0.64	3.12	1.68	0.74	1.00	2.99	0.44	1.79
137人	(1.72)	(2.67)	(1.53)	(0.81)	(2.48)	(1.68)	(0.75)	(0.88)	(2.01)	(0.87)	(1.89)
86~100	1.90	2.54	1.93	0.81	3.17	1.45	0.78	0.84	3.16	0.50	1.54
130人	(1.61)	(2.33)	(1.58)	(0.92)	(2.20)	(1.45)	(0.78)	(0.80)	(2.29)	(1.12)	(1.92)
101~115	2.21	3.33	2.15	0.65	3.33	2.13	0.85	0.92	3.25	0.38	0.85
48人	(1.71)	(2.88)	(1.78)	(0.84)	(2.21)	(1.61)	(0.74)	(0.94)	(1.91)	(0.73)	(1.09)
116以上	1.07	3.07	2.33	0.93	3.73	1.60	0.73	0.73	3.27	0.60	0.60
15人	(0.96)	(1.87)	(1.54)	(1.10)	(2.29)	(1.50)	(0.70)	(1.10)	(2.05)	(1.18)	(0.74)

*上段が平均値、下段が標準偏差である。

*類似は粗点が0点の被験者が9人いたため、FSIQ: ≤ 70 は39人、FSIQ:71~85は132人、FSIQ:86~100は127人であった。

*理解は粗点が0点の被験者が1人いたため、FSIQ:86~100は129人であった。

*逆唱は粗点が0点の被験者が5人いたため、FSIQ: ≤ 70 は39人、FSIQ:71~85は134人、FSIQ:86~100は129人であった。

*語音整列は粗点が0点の被験者が4人いたため、FSIQ:71~85は135人、FSIQ:86~100は128人であった。

点の被験者が5人いたため、FSIQ：≤70群は39人、FSIQ:71～85群は134人、FSIQ:86～100群は129人であった。「語音整列」は粗点が0点の被験者が4人いたため、FSIQ:71～85群は135人、FSIQ:86～100群は128人であった。

FSIQ水準による誤答数の違いを下位検査毎に一要因分散分析を行った。その結果「類似:F(4,356)=1.72,p=.14, $\eta^2=.02$ 」「単語:F(4,365)=1.80,p=.13, $\eta^2=.02$ 」「理解:F(4,364)=0.39,p=.82, $\eta^2<.01$ 」「積木模様:F(4,365)=1.13,p=.34, $\eta^2=.01$ 」「絵の概念:F(4,365)=0.34,p=.85, $\eta^2<.01$ 」「行列推理:F(4,365)=1.83,p=.12, $\eta^2=.02$ 」「順唱:F(4,365)=0.37,p=.83, $\eta^2<.01$ 」「逆唱:F(4,360)=0.74,p=.56, $\eta^2<.01$ 」「語音整列:F(4,361)=1.55,p=.19, $\eta^2=.02$ 」「符号:F(4,365)=0.52,p=.72, $\eta^2=.01$ 」では有意差が見られなかった。唯一「記号探し:F(4,365)=3.80,p<.01, $\eta^2=.04$ 」において5%水準で有意差が見られたため、多重比較を行った。その結果、5%水準の有意差でFSIQ:86～100群がFSIQ:101～115群よりも誤答数が多い結果が得られた(t(365)=3.13,p=.02,d=0.40)。「記号探し」におけるFSIQ:86～100群の誤答数は平均値と標準偏差を足して判断基準を算出したところ4個となり、FSIQ:101～115群では2個となった。なおFSIQ≤70群は4個、FSIQ:71～85群は4個、FSIQ>115群は2個であった。この結果から、FSIQ:100以下の場合は4個以上、FSIQ:101以上の場合は2個が基準値と見なした。

3.4 ISSの大きい児童、ISSの無い人数

各下位検査におけるISSの大きい児童、ISSの無い児童の人数と被験者児童全体に対する割合を表6に示す。

なお年齢毎のISSの平均値に違いが見られた「単語」「理解」「語音整列」は以下のようにISSの大きい基準を変えて人数を算出した。「単語」は5歳から9歳は4個以上、10歳以上は7個以上を基準に人数を算出した。「理解」において、5歳の児童は3個、6歳以上は4個を基準に人数を算出した。「語音整列」では5歳から6歳は4個以上、7歳以上は6個以上を基準に人数を算出した。FSIQ水準では「記号探し」で有意差が見られ、FSIQ:100以下は4個、FSIQ:101以上は2個でISSの基準を変えて人数を算出した。

その結果、ISSが大きい児童の人数は次の通りであった。「類似」57人、「単語」64人、「理解」69人、「積木模様」68人、「絵の概念」65人、「行列推理」42人、「順唱」55人、「逆唱」83人、「語音整列」57人、「符号」43人、「記号探し」53人であった。ISSの無い児童の人数は次の通りであった。「類似」76人、「単語」59人、「理解」73人、「積木模様」193人、「絵の概念」38人、「行列推理」130人、「順唱」151人、「逆唱」135人、「語音整列」59人、「符号」268人、「記号探し」130人であった。

4. 考察

4.1 基礎的統計量から検討したISSの基準

各下位検査の誤答数の平均値、中央値、標準偏差を算出したところ、どの下位検査の誤答数も平均値と中央値の差はないものと考えられた。一方「積木模様」「符号」「記号探し」は平均値よりも標準偏差の値が大きかった。「順唱」「逆唱」においては平均値と標準偏差がほぼ同じ値に位置するものと考えられ、下位検査の誤答数が0個の児童が多い分布と考えられた。しかしこれらの下位検

表6 ISSの大きい児童、ISSの無い児童の人数

ISS	類似	単語	理解	積木模様	絵の概念	行列推理	順唱	逆唱	語音整列	符号	記号探し
無い	76(21%)	59(16%)	73(20%)	193(52%)	38(10%)	130(35%)	151(41%)	135(36%)	59(16%)	268(72%)	130(35%)
大きい	57(15%)	64(17%)	69(19%)	68(18%)	65(18%)	42(11%)	55(15%)	83(22%)	57(15%)	43(12%)	53(14%)

* ()内は全体の児童に対する割合を意味する。

査のISSの大きい群の人数は53人～68人であり、被験者児童全体に対する割合は14%から18%であった。その他の下位検査のISSの大きい群の人数は42人～83人であり、被験者児童全体に対する割合は11%から22%であり、概ね同じ割合と考えられた。以上を踏まえると「積木模様」「順唱」「逆唱」「符号」「記号探し」のISSが大きいと判断する基準値は表3の示す値で良いと考えられる。本研究ではISSが大きいとする基準を平均値よりも1標準偏差大きいものと考えた。ただし誤答数0個に該当する児童が多いと考えられる「積木模様」「行列推理」「符号」「記号探し」の下位検査に関して、ISSが大きい基準値以下でもISSが大きい可能性を考慮に入れた解釈を検討していくことが必要と思われる。本研究では各下位検査の平均値に1標準偏差を加え、小数点が見られた場合には1繰り上げる処理を行ったものの、これらの下位検査では平均値に標準偏差を足した整数の値も基準値として考えた方がよいかもしれない。例えば「行列推理」のISSが大きい基準値は4個である。しか誤答数が4個以上の児童は42人に対し、誤答数が0個の児童は130人とISSの無い児童の方が多し。「行列推理」のISSの平均値(1.62)に1標準偏差(1.58)を足した値は3.20と4よりも3に近い値であり、誤答数が3個の児童でもISSが大きいと判断して解釈することが必要かもしれない。ISSが大きい場合には単独で解釈していくものではなく、仮説を立てるために使用するものと考えれば、厳密に平均値から1標準偏差大きいものとするよりも、幅をもって解釈していくことが良いと考えられる。一方「順唱」「逆唱」は誤答数が0個に該当する児童が多いものの、ISSの大きい基準が2個のため、1個の児童も対象にすると、ISSが大きい児童の方が大半を占める。そのため「順唱」「逆唱」は幅をもって解釈していくよりも、誤答数が2個以上をISSが大きい基準にすべきと考えられる。「積木模様」も誤答数が1個の児童を対象にするとISSが大きい児童が半数を占めるため、誤答数が2個以上を基準にすべきと考えられる。一方、「符号」は誤答数が0個の児童が7割

を占めるため、誤答数が1個の場合にもISSが大きい可能性を考えて検討することが望ましいと考えられる。

またISSが無い群とISSが大きい群が概ね同程度の割合である「類似」「単語」「理解」「絵の概念」「語音整列」に関しては、ISSが無い場合にも統計的には出現率が低い傾向と考えられる。このような下位検査ではISSが無い場合でも検査場面の回答反応を検討することが有益な可能性が考えられ、他の下位検査と得点差が見られない場合や、下位検査得点が平均の水準に位置する場合でも誤答の分析をしていくことが望ましいと考えられる。例えば「単語」において開始問題から2点の正答が連続で得られ、その後誤答が連続で続き中止条件に至った場合、既有知識を的確に説明する力があることにに対し、語彙が不足していたり、難しい問題になるとモチベーションが下がりやすかったりする特性を反映しているかもしれない。このような特性については背景情報や必要に応じて追加の検査などを検討することが必要となるものの、ISSが無い場合にも注目して児童の特性を掘り下げて検討することが臨床上有用と考えられる。

4.2 年齢、FSIQ 毎の違いから検討したISSの基準

各下位検査の誤答数が年齢によって変わるのか検証したところ、「単語」では年齢による誤答数の違いがあると考えられた。有意差の見られなかった年齢群を1つの群にしたところ、5歳から9歳と10歳以上に別れ、ISSの基準値は5歳から9歳が4個以上の誤答数、10歳以上は7個以上の誤答数がISSの大きさの基準と考えられた。「理解」に関しては11歳群と5歳群の誤答数に差があった。5歳群の判断基準値は3個以上であり、11歳群の誤答数の判断基準値は5個以上と考えられた。「理解」の他の年齢群の誤答数の基準値は4個から5個であり、6歳以上のISSの平均値に1標準偏差を加えた結果、6歳以上のISS基準値は4個以上と考えられた。そのため「理解」に関

して、5歳は3個、6歳以上は4個以上がISS基準値と考えられる。「語音整列」では5歳群、6歳群が多く年齢群よりも低い傾向にあった。5歳から6歳を1つの群、7歳以上を1つの群と捉えた。平均値、標準偏差を算出し、それぞれの平均値に1標準偏差を加えたところ、5歳から6歳は4個以上、7歳以上は6個以上がISS基準値と考えられた。FSIQ水準によって誤答数が変わるのか検証したところ、記号探しでは有意差が見られた。FSIQ:100以下の群では、ISSが大きい基準値は4個以上であることに對し、FSIQ:101~115のISSが大きい基準値は2個以上であった。FSIQ>115群でもISSの大きい基準値は2個と考えられた。そのためFSIQ水準が101以上の場合にはISSが大きい基準値は2個と考えられた。

以上の年齢やFSIQによる誤答数の判断基準の違い、幅をもって誤答数の解釈を変えた方がよい下位検査、誤答数が0個の場合も検討に価値がある下位検査を表7に示す。

4.3 本研究の成果と今後の課題

WISC-IVの下位検査内のばらつきに関する基準値を作成した。このような基準値が作成された

ことにより、検査者のISSの大きさを踏まえた質的情報の解釈をしていく判断を支えるものと考えられる。誤答数が多い場合には単独で解釈していくのではなく、複数の情報と統合して解釈していく必要がある。またISSが無い場合にも回答反応を注意深く検討する必要性が考えられた。ISSの大きさは仮説を考えるものであり、今後は仮説から児童の特性を検討する事例の蓄積が期待される。

【文献】

- Courville,T.,Coalson,D,L.,Kaufman,A,S.,& Raiford,S,E. (2016):Does WISC- V Scatter Matter? In Kaufman,A,S.,Raiford,S,E.,&Coalson,D,L.(Eds),Intelligent Testing with The WISC- V .Wiley,NewJersy, pp.209-225.
- Dumont,R.& Willis,J.O.(1995):Intrasubtest Scatter on the WISC- III For Various Clinical Samples vs The Standardization Sample : An Examination of WISC Folkore.Journal of Psychoeducational Assessment,13(3),271-285.
- Elis,Paul,D.(2010):The Essential Guide to Effect Sizes.Cambridge University Press,NewYork.
- Godber,T.,Anderson V., & Bell R.(2000):The

表7 ISSが大きいと判断する誤答数の値,ISSが無い場合でも有益な下位検査

	下位検査	ISS が大きいと判断する誤答数	ISS が大きい可能性がある誤答数	誤答数 0 個の場合でも仮説を立てることが有益
VCI	類似	4 個		○
	単語	5 歳～9 歳:4 個		○
		10 歳以上:7 個		
PRI	理解	4 個(5 歳は 3 個)		○
	積木模様	2 個		○
	絵の概念	6 個		
WMI	行列推理	4 個	3 個	
	順唱	2 個		○
	逆唱	2 個		
PSI	語音整列	6 個 (5～6 歳は 4 個)		
	符号	2 個	1 個	
	記号探し	4 個 (FSIQ:100 よりも 上の場合は 2 個)	3 個 (FSIQ:100 よりも 上の場合は 1 個)	

- measurement and diagnostic utility of intrasubtest scatter in pediatric neuropsychology. *Journal of Clinical Psychology*, 66(1), 101-112.
- Hale, J. B., & Fiorello, C. A. (2004): *School neuropsychology: A practitioner's handbook*. New York : Guilford Press.
- Juni, S., & Trobliger R. (2009): Codification of intratest scatter on the Wechsler Intelligence Scales: Critique and proposed methodology. *Canadian Journal of School Psychology*, 24, 140-157.
- Kaplan, E. (1988): *The process approach to neuropsychological assessment*. *Aphasiology*, 2:3-4, 309-311.
- 松田修 (2015): 日本版 WISC-IV による発達障害のアセスメント—代表的な指標パターンの解釈と事例紹介— 上野一彦, 松田修, 小林玄, 木下智子. 第3章指標 WISC-IV によるアセスメントの手順. 日本文化科学社. pp108-119.
- McCloskey & Maerlender (2005): WISC-IV Integrated. In Prifitera, A., Saklofske, H., Weiss, L. G. [Eds.] *WISC-IV Use and Interpretation. Scientist-Practitioner Perspectives*. Elsevier/Academic Press, San Diego, CA. (上野一彦監訳, (2012): *WISC-IV の臨床的利用と解釈* 日本文化科学社, pp. 125-179.)
- Miller, D. C., & Jones, A. M. (2016): Interpreting the WISC-V from DAN Miller's Integrated school Neuropsychological/Cattell-Horn Carroll Model. In Kaufman, A. S., Raiford, S. E., & Coalson, D. L. (Eds.), *Intelligent Testing with The WISC-V*. Wiley, New Jersey, pp. 459-470.
- 岡田智, 田邊李江, 飯利知恵子, 小林玄, 鳥居深雪 (2015): 日本版 WISC-IV における検査行動アセスメントの意義と実践的課題—検査行動チェックリストの作成と事例による検討—. *子ども発達臨床研究*. 7, 23-35
- 大久保銈, 岡田謙介 (2012): 伝えるための心理統計. 勁草書房
- 竹厚誠, 池田夏葉, 山内裕子他 (2014): Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV) 「絵の抹消」の有用性. *The Japanese Journal of Rehabilitation Medicine*. 51(10), 654-661.

Development of the criteria score intra-subtest scatter(ISS) in WISC-IV

Takumi Sato Satoshi Kanda Kanako Tatehana

Abstract

This study examined the criteria scores of intra-subtest scatter(ISS) in WISC-IV basic subtest. 370 children participated in this study. 341 children had either of diagnosis of ASD, AD/HD, SLD, ID, or DCD, and 29 children didn't have a diagnosis. We calculated the mean value and standard deviation of the number of incorrect answers. We excluded incorrect answers subject to the discontinue rule for each subtest from the number of incorrect answers. If the number of incorrect answers was one standard deviation larger than the mean number of incorrect answers, the ISS was considered to be large. Result showed criteria scores of ISS were following scores. Similarities was 2, Block Design was 4, Picture Concept was 6, Matrix Reasoning was 4, Digit span Forward and Digit Span Reverse were 2, and Coding was 2. Vocabulary, Comprehension, and Letter-Number Sequencing showed to have varying ISS criterion scores depending on age, and Symbol Search showed to have varying criterion scores depending on FSIQ level.

Key words: WISC-IV, Intra-subtest scatter, assessment
