

厚生労働科学研究費補助金（障害者政策総合研究事業）  
分担研究報告書

知能および適応行動の標準化検査の整合性に関する心理測定的検証

分担研究者 伊藤大幸 お茶の水女子大学基幹研究院人間科学系

### 研究要旨

国際的診断基準では、知的障害の診断と重症度の判定を知能および適応行動に関する標準化検査の結果に基づいて行うことが求められている。しかし、標準化検査であっても測定の精度（信頼性・妥当性）は検査によって異なり、必ずしも相互に整合的な検査結果をもたらすとは限らない。そこで本研究では、療育手帳判定における知能および適応行動の評価に関して、2つの観点から心理測定的検証を行った。第1に、幼児期から成人期までの療育手帳保有者を対象として3つの知能検査と2つの適応行動尺度を実施し、得点差や相関の観点から検査結果の整合性を検証した。第2に、成人の一般サンプルおよび知的障害者のデータに基づいて、知能検査と適応行動尺度の検査結果の合成による判定精度の向上について検証した。これらの検証の結果、検査間での測定結果の系統的な乖離とランダムなブレは無視できないほど大きいことから、いずれの検査も単独で療育手帳の交付や重症度の判定に用いることは適切でなく、知能検査と適応行動尺度の結果を組み合わせることで評価を行うことの重要性が示された。

### A. 研究目的

療育手帳制度は、1973年の厚生事務次官通知「療育手帳制度について」に基づく制度であり、知的障害児者に対する主要な施策の一つであるが、制度開始以来、統一的な規定は設けられておらず個々の自治体が独自に規定・運用を行っている。精神医学領域では、1980年代以降、DSM（APA, 2013）やICD（WHO, 2015）などの操作的診断基準が整備されてきており、知的障害の診断と重症度の判定は、知能および適応行動に関する標準化検査の結果に基づいて行うことが求められている。

こうした流れに一致して、各自治体の療育手帳判定においても標準化された知能検査や発達検査が用いられているが、複数の実態調査（柴田, 2004; 吉村ら, 2019）や本研究班の小林（2021）の分担報告が明らかにしているように、使用される検査や判定の基準は自治体によって異なっている。標準化された検査であっても、標準化の手続きがいつ、どのように行われたのか、測定の精度（信頼性・妥当性）がどのように検証されているのかは検査によって様々であり、必ずしも全ての検査が同等の心理測定的な質を備えているとは限らない。療育手帳

の交付や重症度判定の基準の統一化を図るためには、まず複数の標準化検査の間で、検査結果にどの程度の整合性があるのかを検証することが不可欠である。

また、知能と適応行動という2つの指標をどのように組み合わせて判定を行うかという問題についても検討の必要がある。DSM-5 (APA, 2013) や ICD-11 (WHO, 2018) では、知能だけでなく、適応行動の検査結果を考慮して知的障害の診断を行うことが求められるようになった。これは、知能検査によって測定されたIQが、実際の生活場面における困難さを推し量る上で必ずしも十分な指標ではないことを踏まえたものである (APA, 2013)。検査の測定値には検査の信頼性・妥当性の範囲内でランダムまたは系統的な測定誤差が含まれるという統計学的事実からも、2つの独立した検査の結果を総合して判断を行うことには合理性があると言える。しかし、実際の運用において、知能と適応行動という2つの指標をどのように組み合わせて診断を行うかについて、必ずしも一致した指針は示されておらず、療育手帳判定においても、自治体間で明確な方針は共有されていない。

こうした現状を踏まえ、本稿では、療育手帳判定における知能および適応行動の評価に関して、2つの観点から多面的な心理測定学的検証を行った。第1に、幼児期から成人期までの療育手帳保有者を対象として3つの知能検査と2つの適応行動尺度を実施し、得点差や相関の観点から検査結果の整合性を検証した。第2に、成人の一般サンプルおよび知的障害者のデータに基づいて、知能検査と適応行動尺度の検査結果の組み合わせに基づく療育手帳判定の方法につい

て検証した。

## B. 研究方法

研究1：療育手帳保有する幼児から成人を対象として代表的な知能検査および適応行動尺度を並行実施した。知能検査として、ウェクスラー式検査 (WPPSI-III、WISC-IV、WAIS-IV；以下ウェクスラー式)、田中ビネー検査 (以下ビネー式)、新版K式 (以下K式)、適応行動尺度として Vineland-II 適応行動尺度 (以下 Vineland) および S-M 社会生活能力検査 (以下 S-M) を実施した。ただし、幼児期 (3~5歳) の参加者には5つの検査全て、児童期 (6~15歳) の参加者にはK式を除く4つの検査、成人期 (16歳以上) の参加者にはK式とS-Mを除く3つの検査を実施した。ウェクスラー式検査は、幼児期 (3~5歳) の参加者には WPPSI-III、児童期 (6~15歳) の参加者には WISC-IV、成人期 (16歳以上) の参加者には WAIS-IV を使用した。ウェクスラー式検査および Vineland では偏差指数、ビネー式、K式、S-M は比例指数の標準得点が算出された。98名が調査に参加したが、一部の検査が実施できなかった参加者 (26名) を除き、72名のデータを分析に使用した。参加者、調査内容、手続きの詳細については、本研究班の村山・浜田 (2022) の分担研究報告書を参照されたい。

研究2：定型発達者と知的障害者のデータを使用して、知能検査と適応行動尺度の結果を合成することによる判定精度の向上について検証する。このデータは、厚生労働省令和2年度社会福祉推進事業「日常生活支援住居施設の対象者選定のためのシステムに関する調査研究事業」(代表者：辻井正

Table 1 年齢区分ごとの各検査の標準得点の平均値および標準偏差

	幼児期 (3-5歳) (n=28)		児童期 (6-15歳) (n=27)		成人期 (16歳以上) (n=17)	
	M	SD	M	SD	M	SD
ウェクスラー	55.36	11.32	57.78	17.36	51.41	10.76
ビネー	56.68	23.34	52.70	26.18	32.12	14.15
K式	56.32	20.45				
Vineland	65.25	15.22	47.22	17.46	28.47	14.73
S-M	55.52	19.69	49.99	20.96		

次)における調査により、一般サンプル 418 名(男性 208 名、女性 211 名;平均年齢 49.8 歳)および知的障害者 33 名(男性 20 名、女性 13 名;平均年齢 49.1 歳)から得られたものである。知能および適応行動の測定には、同事業によって開発された ABIT (Adaptive Behavior and Intelligence Test) を用いた。ABIT は、タブレット端末を用いた包括的アセスメントツールであり、知能については対象者が 3 つの課題に取り組む形式、適応行動については対象者をよく知る同居者・介護者が 35 の項目に評定する形式で測定された。WAIS-IV および Vineland-

II との相関や複数の臨床群の識別精度に基づいて妥当性が検証されている。本研究の手続きは、中京大学現代社会学部倫理審査委員会の審査と承認を得た。

### C. 研究結果

#### 1. 知能および適応行動の標準化検査の整合性に関する検証

##### 1) 検査間の得点差と相関

各検査の標準得点の年齢区分ごとの記述統計量を Table 1、各検査間の相関係数と得点差の平均値および 90%信頼区間を Table 2 に示した。幼児期においては、Vineland の

Table 2 各検査間の相関係数と得点差の平均値および 90%信頼区間

		幼児期 (3-5歳) (n=28)			児童期 (6-15歳) (n=27)			成人期 (16歳以上) (n=17)		
		r	得点差		r	得点差		r	得点差	
			M	90% CI		M	90% CI		M	90% CI
ウェクスラー	ビネー	.64 ***	-1.32	± 30.16	.86 ***	5.07	± 23.52	.81 ***	19.29 ***	± 13.56
ウェクスラー	K式	.68 ***	-0.96	± 25.02						
ウェクスラー	Vineland	.59 ***	-9.89 ***	± 20.51	.39 *	10.56 **	± 31.61	.54 *	22.94 ***	± 20.89
ウェクスラー	S-M	.56 **	-0.16	± 26.78	.64 ***	7.79 *	± 27.41			
ビネー	K式	.88 ***	0.36	± 18.21						
ビネー	Vineland	.74 ***	-8.57 **	± 26.06	.64 ***	5.48	± 33.11	.77 ***	3.65	± 16.18
ビネー	S-M	.51 **	1.16	± 35.32	.70 ***	2.72	± 31.06			
K式	Vineland	.74 ***	-8.93 **	± 22.76						
K式	S-M	.61 ***	0.80	± 29.28						
Vineland	S-M	.78 ***	9.73 ***	± 20.42	.69 ***	-2.77	± 25.31			

\*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

みが他の検査よりも有意に高い得点を示し、その得点差は約 9~10 点であった。児童期では、ウェクスラー式の得点が 2つの適応行動尺度より有意に高く、約 8~11 点の差が見られた。成人期では、ウェクスラー式がビネー式および Vineland より有意に高い得点を示し、得点差は約 19~23 点に達した。以上の結果から、幼児期では Vineland、児童期や成人期ではウェクスラー式が他の検査よりも系統的に高い得点を示す傾向があることが示された。

併存的妥当性の根拠となる検査間の相関

係数はいずれも有意な値を示したが、組み合わせによって値の違いが見られた。3つの知能検査間の相関は、幼児期ではビネー式と K 式が .88 という十分な値を示したのに対し、ウェクスラー式と他の 2 検査の相関は .64 および .68 とやや低い値を示した。児童期および成人期におけるウェクスラー式とビネー式の相関は、それぞれ .86 および .81 という十分な値を示した。2つの適応行動尺度間の相関は、幼児期で .78、児童期で .69 であり、児童期はやや低い水準にあった。知能検査と適応行動尺度の相関を見る

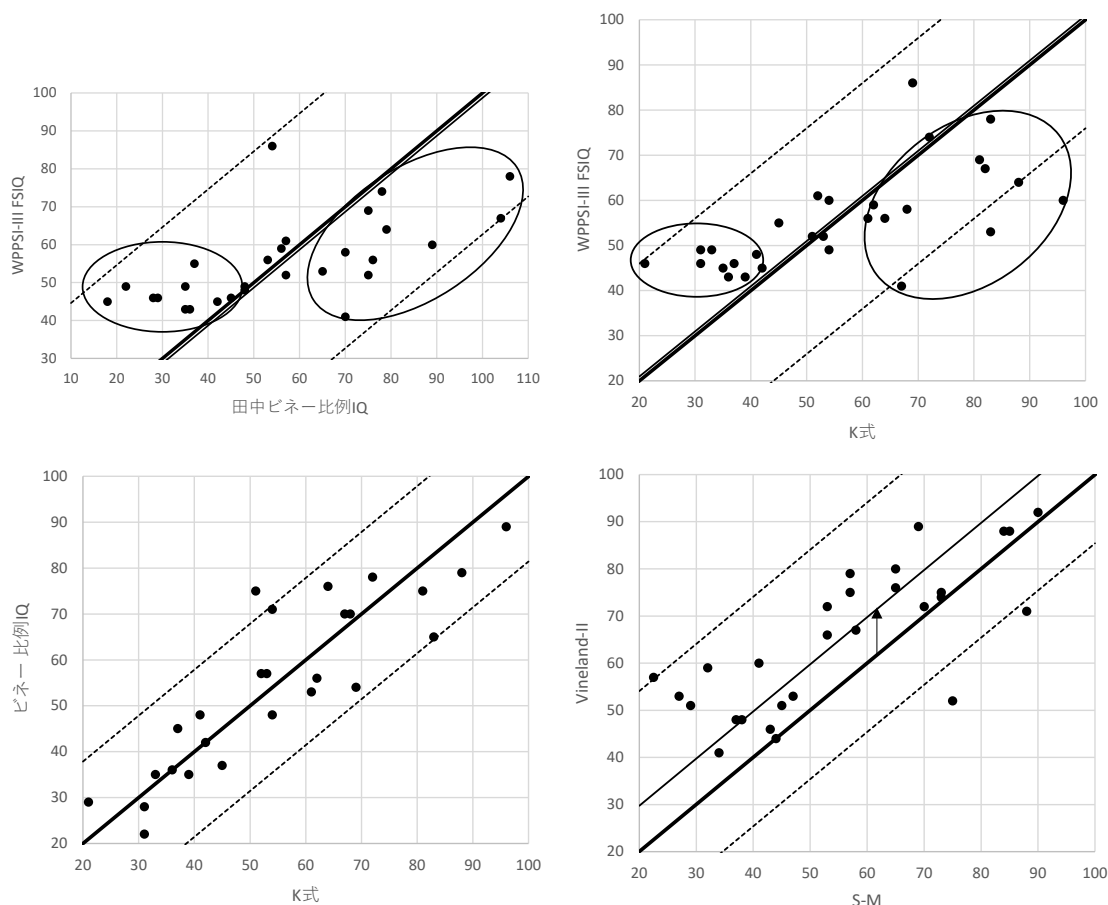


Figure 1 幼児期における各検査の標準得点間の散布図

注：知能検査間の散布図および適応行動尺度間の散布図のみを示した。太い実線は  $y=x$  のグラフ、細かい実線は系統的な得点差を補正した予測値のグラフ ( $y=x+$ 得点差の平均値)、破線は 90%信頼区間を示す。楕円は予測値からの系統的な逸脱を示すデータ群、矢印は検査間の系統的な得点差（統計的に有意なもののみ）を表している。

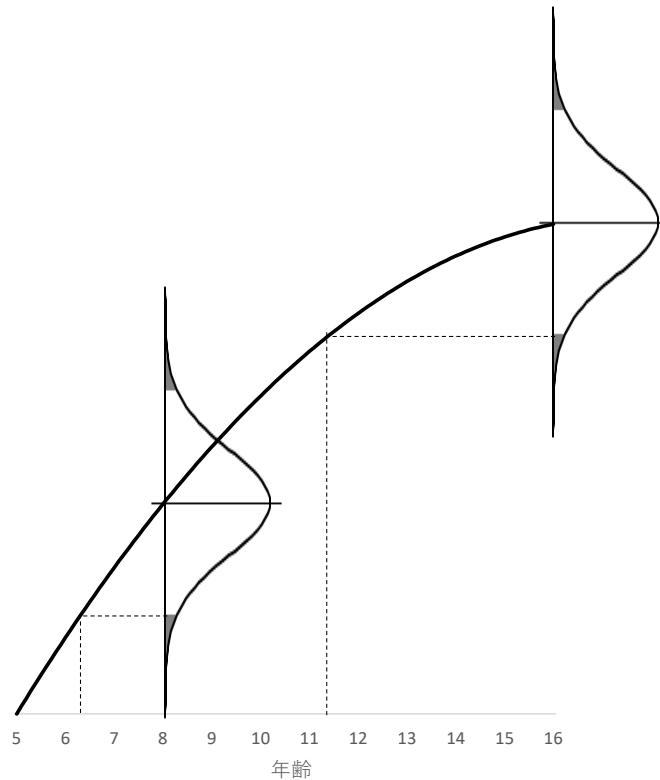


Figure 2 偏差指数と比例指数の関係に関する模式図

注：グラフは年齢にともなう粗点の推移（仮想）を示す。8歳時点と16歳時点において、それぞれ偏差指数で70（平均-2SD）に値する粗点が観測されたとき、年齢による粗点の変化の傾きが大きい8歳時点における精神年齢は6.2歳（生活年齢-1.8歳）であるのに対し、傾きが小さい16歳時点では精神年齢が11.3歳（生活年齢-4.7歳）となり、比例指数は8歳時点で78、16歳時点で71となる。

と、Vinelandとの相関はビネー式(.64~.77)やK式(.74)がウェクスラー式(.39~.59)よりも全般的に高い値を示したが、S-Mとの相関はいずれの検査もおおむね同程度の値(.51~.70)を示した。以上より、併存的妥当性の観点で、3つの知能検査の中ではウェクスラー式が他の2検査よりもやや低い水準にあることが示唆された。一方、適応行動尺度であるVinelandとS-Mの間では明瞭な結果の違いは見られなかった。

Figure 1に、幼児期における各検査の標準得点間の散布図を示した。ウェクスラー式(WPPSI-III)とビネー式およびK式の

散布図(上部の2つ)を見ると、平均値としての系統的な得点差は見られないものの、個々の測定値のレベルでは他方の検査結果に基づく予測値からの系統的な逸脱(楕円で示した部分)が生じていることが見て取れる。第1に、横軸(ビネー式およびK式)の値が40を下回る範囲では、ウェクスラー式の測定値がビネー式よりも系統的に高い値を示している。これはビネー式やK式の標準得点が40未満の値を取りうるのに対し、ウェクスラー式では標準得点の下限が40に設定されているため、フロア効果により得点を実際よりも過大に推定されている

ことを反映している。第2に、横軸の値が60を超える範囲でウェクスラー式の測定値がビネー式よりも系統的に低い値を示している。これはウェクスラー式の標準得点が偏差指数（各年齢帯における相対的位置を表す指数）であるのに対し、ビネー式やK式の標準得点が比例指数（精神年齢と生活年齢の比を取った指数）であることによるものと考えられる。Figure 2 に模式図を示したように、一般的に年齢上昇にともなう知能検査の粗点の傾きは幼少期で最も大きく、成人期にかけて減少していくことから、偏差指数が同一であった場合、低年齢ほど比例指数が低く推定されやすくなる。つまり、低得点域での予測値からの逸脱がウェクスラー式の得点の下限の問題によるものに対して、高得点域での逸脱はビネー式やK式が採用する比例指数の問題によるものであると考えられる。この解釈を支持するように、得点の下限や比例指数の採用に

おいて共通するビネー式と K 式の散布図（左下）では予測値からの系統的な逸脱が見られず、信頼区間の幅も大幅に縮小している。一方、適応行動尺度間の散布図では、Vineland が S-M よりも系統的に高い測定値を示しているが、個々の測定値のレベルでは予測値からの系統的な逸脱は見られない。

次に、児童期の散布図を Figure 3 に示す。ウェクスラー式とビネー式の散布図では、幼児期と同様に、系統的な得点差は見られないものの、低得点域でウェクスラー式のフロア効果により予測値と実測値の逸脱が生じている。一方、高得点域の系統的な逸脱は解消されていることが見て取れ、中間域にあたるこの年代では偏差指数と比例指数の乖離が生じにくくなると考えられる。Vineland と S-M の間では系統的な得点差がなくなり、測定値レベルでの系統的な逸脱も見られない。ただし、得点間の相関がや

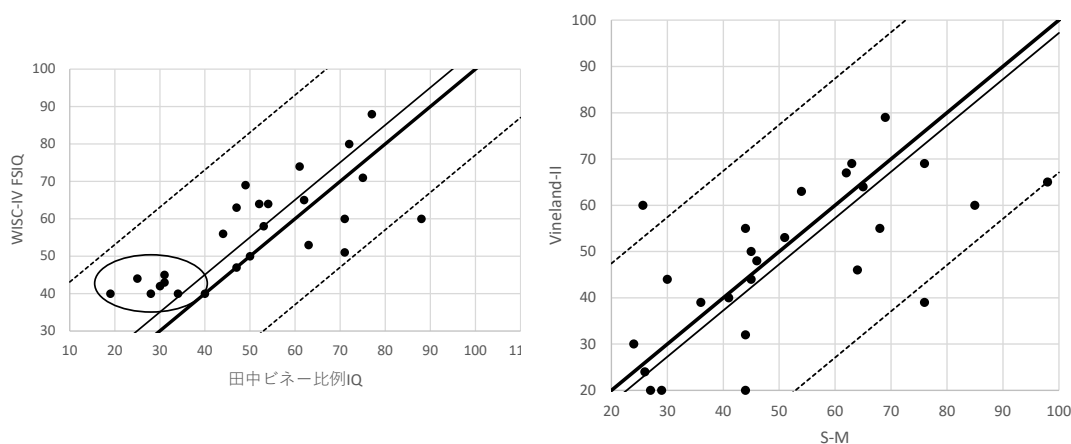


Figure 3 児童期における各検査の標準得点間の散布図

注：知能検査間の散布図および適応行動尺度間の散布図のみを示した。太い実線は  $y=x$  のグラフ、細かい実線は系統的な得点差を補正した予測値のグラフ ( $y=x+$ 得点差の平均値)、破線は 90%信頼区間を示す。楕円は予測値からの系統的な逸脱を示すデータ群を表している。

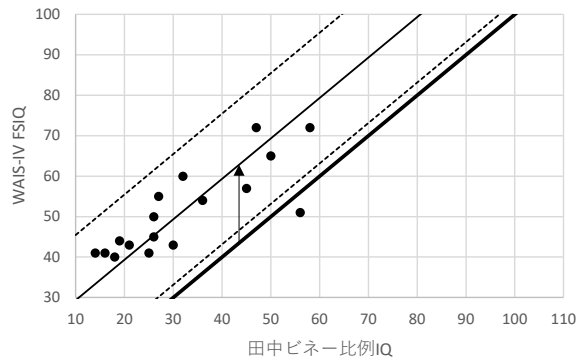


Figure 4 成人期における各検査の標準得点間の散布図

注：知能検査間の散布図のみを示した。太い実線は  $y=x$  のグラフ、細い実線は系統的な得点差を補正した予測値のグラフ ( $y=x+$ 得点差の平均値)、破線は 90%信頼区間を示す。矢印は検査間の系統的な得点差（統計的に有意なもののみ）を表している。

や低いため ( $r=.69$ )、90%信頼区間の幅は比較的広がっている ( $\pm 25.31$ )。

成人期におけるウェクスラー式とビネー式の散布図 (Figure 4) では、平均値のレベルで顕著な得点差 (19.29) があることが見て取れる。これはウェクスラー式のフロア効果により低得点域の得点が過大推定されていることに加え、年齢の上昇にともない、比例指数が偏差指数に対して過小に推定されるようになったことが組み合わさった結果と考えられる。

以上の結果から、第1に、知能検査間の整合性については、得点の下限や得点算出方式の違いにより、ウェクスラー式とビネー式およびK式の得点に無視できない程度の結果の乖離が生じることが示された。また、その乖離の表れ方は得点域や年代によって顕著に異なる複雑な様相を呈しており、単純な換算式の設定などにより解消することは困難である。第2に、適応行動尺度間の整合性については、幼児期では Vineland が S-M よりも系統的に高い得点を示した一方、児童期では系統的な得点差が消失するもの

の、得点間の相関はやや低く、検査間のランダムな得点のブレは小さくないことが明らかになった。つまり、知能検査・適応行動尺度のいずれについても、系統的な得点の乖離 (妥当性の問題) とランダムな得点のブレ (信頼性の問題) が検査間の結果の整合性を毀損しており、いずれの検査を使用するとしても、単一の検査の結果にのみ基づいて機械的に診断や判定を行うことは適切でないと結論づけられる。実際、検査間の得点差の 90%信頼区間の幅は、大きいところで 30 (2SD) にまで達している (例えば幼児期のウェクスラー式とビネー式)。これは、例えば WPPSI で 50 という得点が得られたとしても、田中ビネーの測定値は 90%の確率で 20~80 (つまり最重度から健常まで) の範囲でばらつくことを意味しており (Figure 1 を見れば、これが誇張でないことは明白である)、個人の診断や判定のための情報としては、ほとんど有用性がないと言っても過言ではない。

## 2) 知能検査と適応行動尺度の得点の合成

上記のような測定の信頼性・妥当性の問題に対処する上で、簡便かつ有効な方法として得点の合成がある。信頼性に関して言えば、複数の尺度の得点を合成した場合の信頼性は以下のように表すことができる (Wang & Stanley, 1970)。

$$r = \frac{\frac{r_1 + r_2}{2} + r_{1,2}}{1 + r_{1,2}}$$

ここで $r_1$ は検査1の信頼性係数、 $r_2$ は検査2の信頼性係数、 $r_{1,2}$ は2つの検査の相関係数である。例えば、信頼性係数が.80の検査が2つあり、その相関が.60である場合、両者の標準得点を合成(加算または平均)した値の信頼性係数は.91に上昇し、測定値の90%信頼区間は0.74SDから0.50SDまで縮小する。妥当性に関して言えば、各検査が固有の妥当性の問題を抱えていたとしても、それが互いに異なる性質の問題であれば、複数の検査結果を合成することで結果の偏りを相殺することができる。また、内容的妥当性の観点から考えても、知的障害の診断には、互いに重なりを持ちながら独立した構

成概念である知能と適応行動の評価が必要となるため、知能検査と適応行動尺度の得点を合成することで、より領域代表性の高いアセスメントが可能となる。

本研究では3つの知能検査と2つの適応行動尺度を実施したため、知能検査と適応行動尺度の合成得点として3×2の6つの得点を得ることができる。これら6つの合成得点(知能検査と適応行動尺度の標準得点を単純平均した値)の記述統計量をTable 3、相互の相関係数および得点差の平均値と90%信頼区間をTable 4に示す。得点差の平均値に関しては、個別の検査の結果(Table 2)では最大で10程度の差が見られたが、合成得点(Table 4)では5程度までの差に留まっている。相関係数は、個別の検査(Table 2)では知能検査間で.64~.88、適応行動尺度間で.69~.78であり、(フィッシャー変換の上で)平均すると.77であったが、合成得点(Table 4)では.74~.93であり、平均は.87であった。90%信頼区間の幅は、個別の検査(Table 2)では知能検査間で18.21~30.16、適応行動尺度間で20.42~25.31であり、平均は23.77であったが、合成得点では11.40~19.99であり、平均は16.26と、約3分の2程度まで縮小した。こ

Table 3 年齢区分ごとの知能検査・適応行動尺度の合成得点の平均値および標準偏差

	幼児期 (3-5歳) (n=28)		児童期 (6-15歳) (n=27)		成人期 (16歳以上) (n=17)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
ウェクスラー & Vineland	60.30	11.88	52.50	14.52	39.94	11.23
ウェクスラー & S-M	55.44	13.84	53.88	17.35		
ビネー & Vineland	60.96	18.04	49.96	19.84	30.29	13.58
ビネー & S-M	56.10	18.73	51.35	21.75		
K式 & Vineland	60.79	16.64				
K式 & S-M	55.92	17.99				

Table 4 知能検査・適応行動尺度の合成得点間の相関係数と得点差の平均値および90%信頼区間

		幼児期 (3-5歳) (n=28)			児童期 (6-15歳) (n=27)		
		r	得点差		r	得点差	
			M	90% CI		M	90% CI
ウェクスラー & Vineland	ビネー & S-M	.89 ***	4.21 *	± 16.24	.93 ***	3.92	± 18.10
ウェクスラー & S-M	ビネー & Vineland	.74 ***	-5.53 *	± 19.99	.83 ***	1.15	± 16.41
ウェクスラー & Vineland	K式 & S-M	.87 ***	4.38 *	± 15.81			
ウェクスラー & S-M	K式 & Vineland	.80 ***	-5.35 **	± 16.48			
ビネー & Vineland	K式 & S-M	.86 ***	5.05 **	± 15.63			
ビネー & S-M	K式 & Vineland	.93 ***	-4.69 **	± 11.40			

\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$

注：同一の検査を含むペアは除外した。

これは上記の合成得点の信頼性の公式により導かれた結果とほぼ一致している。

以上の結果から、知能検査と適応行動尺度の標準得点を単純平均するという簡便な処理だけで、測定の信頼性・妥当性を大幅に向上しうることが示された。それでも±1SD程度の得点のブレは生じるため、1度の検査結果で機械的・決定的な判定を行うことには慎重である必要があるが、単一の検査の結果のみを用いるよりは幾分精度の高い判定を行うことが可能となる。

## 2. 知能と適応行動の組み合わせに基づく判定方法に関する検討

### 1) 背景と目的

前項の検証により、個々の知能検査や適応行動尺度の検査結果は、相互に整合性が低く、いずれの検査を用いるにしても、単一の検査結果にのみ基づいて判定を行うことは適切でないことが明らかとなった。この問題に対して、知能検査と適応行動尺度の得点を合成することで、一定程度の改善を図ることが可能であることが示された。しかし、この方法が療育手帳判定という目的

においてどの程度の有効性を発揮するかを評価するためには、複数の得点間の整合性（併存的妥当性）だけではなく、定型発達者と知的障害者の識別の精度（識別妥当性）に与える影響について検証する必要がある。

そこで、本項では一般母集団と知的障害者から得られた知能検査および適応行動尺度のデータを用いて、知能と適応行動をどのように組み合わせる判断を行うことが適切なのか、統計学的観点から検証を行う。ただし、他の事業で収集された成人のデータを用いた予備的検証であり、確定的な結論を得るためのものではなく、今後の研究の方向性を検討するための材料を提供することを目的としたものである。

### 2) 散布図での視覚的検討

Figure 5 に一般群と知的障害群における知能と適応行動の散布図を示す。いずれの得点も一般母集団の平均が100、標準偏差が15となるように標準化されている。まずはこの散布図において、一般群と知的障害群を最も精度よく判別する方法はどのようなものであるかを考えてみたい。

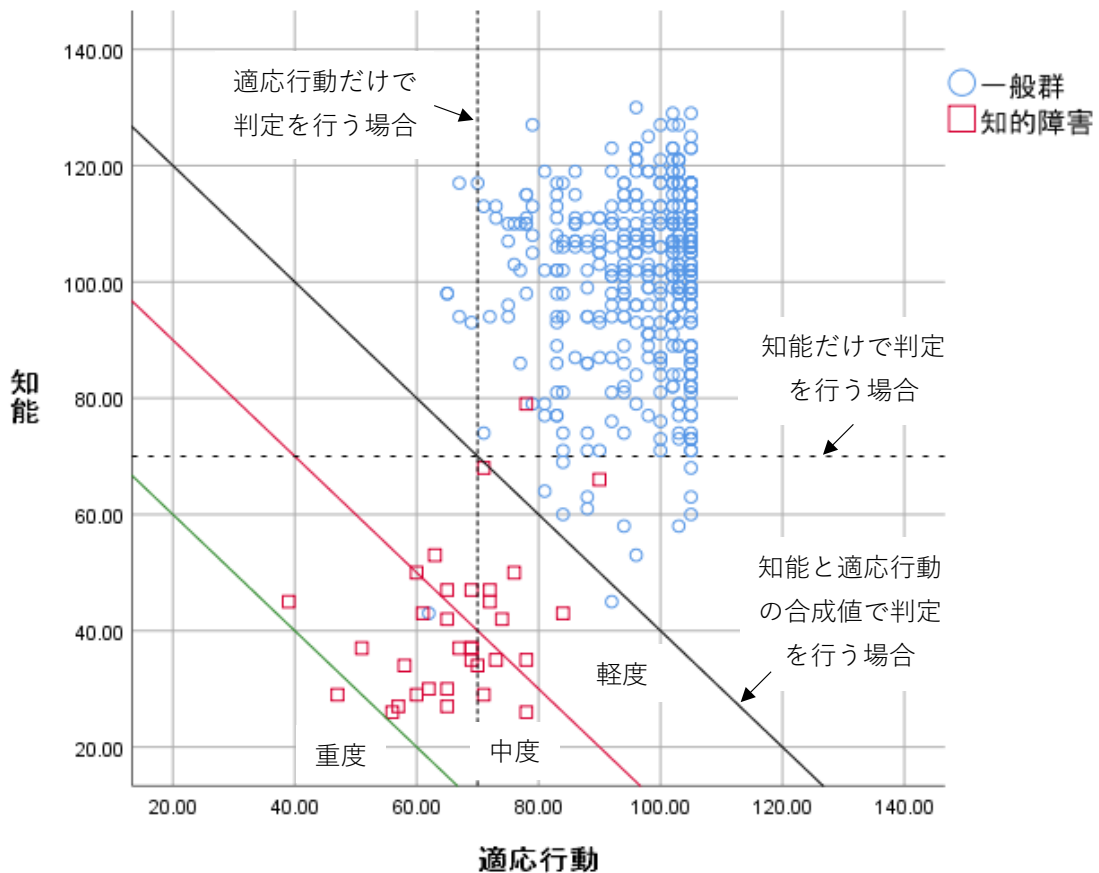


Figure 5 一般群と知的障害群における知能と適応行動の散布図

本研究班の小林（2021）の分担報告によれば、療育手帳判定において、標準化された検査による適応行動のアセスメントを行わず、もっぱら知能検査によるIQによって判定を行っている自治体も多いようである。知能だけで判定を行うということは、図中の縦軸が70の位置に水平方向に引かれた破線で判定を行うことを意味する。この場合、知的障害群の大部分が手帳交付の対象となるが、一般群の一部も対象となってしまうことが見て取れる。一方、適応行動だけで判定を行う場合、横軸が70の位置に垂直方向に引かれた破線で判定を行うことになる。この場合は、一般群の大部分は交付の対象とならないが、知的障害群の多くも交付

の対象から外れてしまう形になる。

これに対し、知能と適応行動の合成値で判定を行うということは、この散布図に斜め方向の直線を引いて判定を行うことを意味する。2つの指標の重みづけの方法によって直線の傾きは変化するものの、ここでは両者を1:1で合成した値（両者の単純平均値）を用いて判定を行う場合の直線を図示した。この場合、水平方向や垂直方向の直線で区切った場合よりも、一般群と知的障害群を精度よく判別できていることが見て取れる。実際、散布図上での一般群の分布位置と知的障害群の分布位置は縦や横ではなく斜めにずれているため、両者ができる限り重ならないように分けるためには、斜め

方向の直線を引くのが合理的であることは明らかである。また、この直線を下方に30ずつ平行移動していくことで、軽度、中度、重度の重症度判定の基準となる境界線を得ることも可能である。

### 3) 判別分析

次に、より定量的なエビデンスを得るため、判別分析による検証を行った。判別分析は、複数の量的な独立変数(ここでは知能と適応行動)により、質的な従属変数のカテゴリ(ここでは一般群と知的障害群)を最も精度よく判別する関数を得るための分析である。独立変数として知能のみを含めた場合、適応行動のみを含めた場合、知能と適応行動の両方を含めた場合の結果を Table 5 に示した。仮に知能か適応行動のうち一方の変数のみで十分な精度の判別が可能であり、他方の変数が独立の貢献を果たさないとすれば、知能と適応行動の両方を分析に含めたとき、いずれかの判別係数は0に近い値を示すはずである。しかし、実際にはいずれも.50以上の判別係数を示しており、両者がそれぞれ一般群と知的障害群の判別に独立の貢献を果たしていることがわかる。また、判別精度の指標である正準相関係数は、知能のみ、適応行動のみを含めた分析よりも高い値を示している。

これらの結果は、知的障害の診断において、知能や適応行動を単独で用いるよりも、知能と適応行動の合成値を用いることが判定の精度を高めることを示唆している。知能と適応行動は相互に独立した構成概念であることに加え、その測定の方法にも違いがある。つまり、知能検査では課題に対する対象者の反応を記録するのに対し、適応行

動尺度では同居者・介護者に普段の対象者の様子について評定を求める。このように、内容と測定方法の両面において異なる指標を組み合わせて用いることで、対象者の実像をより正確に捉え、知的障害者の判別の精度を高めることが可能になっていると考えられる。

判別係数は、個々の変数にどの程度の重みづけをして合成すると判別精度を最大化できるかを表しており、ここでは両者の係数に大きな違いが見られないことから、2)の分析のように、両者の単純平均を用いる形でもおおむね遜色のない判別精度を発揮することができると考えられる。これは実際の運用の利便性を考えたときには、都合の良い性質である(実際、多くの心理尺度では、こうした性質を利用して各項目得点の単純合計を尺度の得点として用いている)。ただし、今回の検討は成人のデータに基づくものであり、実際の運用に適用できる知見を得るためには、乳幼児期から児童・青年期に至るまでの各年齢段階のデータを収集し、判別係数の推定を行うことが必要である。

Table 5 一般群と知的障害群の判別分析

	判別係数		
	知能のみ	適応行動のみ	知能&適応行動
知能	1.000		.770
適応行動		1.000	.558
正準相関	.725	.632	.785

### D. 結論

本研究では、療育手帳判定における知能および適応行動の評価に関して、2つの観

点から心理測定的検証を行った。第1に、幼児期から成人期までの療育手帳保有者を対象として3つの知能検査と2つの適応行動尺度を実施し、得点差や相関の観点から検査結果の整合性を検証した。その結果、検査間での測定結果の系統的な乖離とランダムなブレは無視できないほど大きいことが明らかとなった。系統的な乖離は知能検査間でより顕著であり、その主な原因として、ウェクスラー式知能検査においてIQの下限が40に設定されていることによるフロア効果、田中ビネー検査やK式検査では国際的な診断基準で求められる偏差指数とは性質の異なる比例指数が採用されていることが関係していると考えられた。適応行動尺度間では、幼児期に系統的な平均値の差が見られ、児童期においてはランダムな得点のブレが大きいことが示された。以上より、いずれの検査を用いるにしても、知的障害の有無や重症度について、単一の検査結果により機械的・決定的に判定することは適切でないことが確認された。一方、こうした検査結果のバラつきは、知能検査と適応行動尺度の結果を合成(単純平均)することによって、3分の2程度にまで低減されることが明らかとなった。

第2に、上記のような知能検査と適応行動尺度の検査結果を合成するという手続きによって、知的障害の判定精度がどの程度向上するかについて検証した。成人の一般サンプルおよび知的障害者のデータに基づく検証により、知能、適応行動をそれぞれ単独で用いるよりも、両者の合成値を用いることが知的障害者の判別の精度を高めることが示された。この方法は、手帳交付の判定だけでなく、軽度、中度、重度の重症度の判

定においても有効性を発揮すると考えられる。こうした結果に基づけば、現在、多くの自治体で標準化尺度によって適応行動の評価がなされていないことは重大な問題であり、今後、適応行動尺度の普及を図っていくことが必要であると考えられる。

以上の検討により、いずれの検査も単独で療育手帳の交付や重症度の判定に用いることは適切でなく、少なくとも知能検査と適応行動尺度の結果を組み合わせる評価を行うことの重要性が示された。しかし、療育手帳判定にあたる児童相談所の職員が虐待対応など多くの業務を抱える現状で、検査の負担を増やす方向の提言が受け入れられる可能性は高くない。この問題の根本的な解決には、知的障害の有無と重症度の判定という用途に関して十分な心理測定的性質を持ちつつも、所要時間を大幅に抑えられ、かつ、実施に高度な専門知識が求められないアセスメントツールの開発が必要であると考えられる。我々が並行して進めてきた社会推進福祉事業「日常生活支援住居施設の対象者選定のためのシステムに関する調査研究事業」(代表者：辻井正次)では、成人を対象とした簡便かつ高精度の包括的アセスメントツールであるABITを開発しており、今後、こうしたツールを全年齢に適用できるよう拡張を図る必要がある。

## E. 健康危険情報

該当なし

## F. 研究発表

1. 論文発表 なし
2. 学会発表 なし

## G. 知的財産権の出願・登録状況

該当なし

## H. 引用文献

American Psychiatric Association. (2013).

Diagnostic and statistical manual of mental disorders, fifth edition. American Psychiatric Association, Washington D.C.

小林真理子. (2021). 児童相談所および知的障害者更生相談所における療育手帳の判定基準の統一化に向けての課題の整理. 辻井正次. 令和 2 年度厚生労働科学研究費補助金「療育手帳に係る統一的な判定基準の検討ならびに児童相談所等における適切な判定業務を推進させるための研究」報告書.

柴田長生. (2004). 知的障害と発達診断. 発達, 99, 44-51

内山登紀夫. (2021). 現在の知的障害に関する国際的な診断基準と、最近の知的障害概念の検討. 辻井正次. 令和 2 年度厚生労働科学研究費補助金「療育手帳に係る統一的な判定基準の検討ならびに児童相談所等における適切な判定業務を推進させるための研究」報告書.

World Health Organization. (2018). International statistical classification of diseases for mortality and morbidity statistics (11th Revision). <https://icd.who.int/browse11/l-m/en>

吉村拓馬, 大西紀子, 恵良美津子, 松田裕之, 小橋川晶子, 広瀬宏之, & 大六一志. (2019). 療育手帳判定における知能検査・発達検査に関する調査. *LD 研究*, 28(1), 144-153.