

護などについて説明を行った。

## 5. VABS-II の尺度構成

VABS-II の尺度構成を Table 2 に示す。VABS-II は全 435 項目からなり、大きく適応行動尺度と不適応行動尺度の 2 つに分かれている。適応行動尺度は、4 つの領域から構成され、それぞれに 2 つから 3 つの下位領域が存在する。不適応行動尺度は、不適応内在化、不適応外在化、不適応その他、重要事項の 4 つの下位尺度からなるが、数量的評価には前三者が用いられ、重要事項については個別的な評価のみを行う。また、不適応その他については、下位尺度単位での評価は行わず、不適応行動の総得点を算出する際にのみ用いられる。

## 6. スコアリング

VABS-II の各項目は、基本的に 2、1、0 の 3 段階で評価される。2 点は対象者が手助けなしにその行動を習慣的に行っている場合に与えられる。1 点はその行動の遂行に手助けが必要か、または時々行われている場合に与えられる。0 点は、対象者がその行動を滅多に行わないか、または行われていないことを意味する。

VABS-II の面接調査フォームには、面接の負担を軽くするため、以下のようなルール（以後、打ち切りルールと記載）が存在する。VABS-II の項目は、下位領域ごとに、最も初歩的な適応行動（低年齢の子どもでも可能な行動）に関する項目から徐々に難しい行動に関する項目になるよう配置されている。各下位領域の面接は、1 番目の項目から始まるのではなく、対象者の年齢に応じて異なる項目から開始される。その開

始項目について「2」の評価がついた場合、評価は後の項目に進むが、「1」や「0」の評価がついた場合は逆方向に戻って評価が行われる。逆方向に評価を行う場合、連続して 4 回、「2」の評価がついた時点でそれより前に位置する項目を全て「2」の評価とし、今度は開始項目から順方向に評価を進めていく。その後、「0」の評価が連続して 4 回続いた時点で評価を打ち切り、その後の項目を全て「0」の評価とする。VABS-II は、このような打ち切りルールを設定することで、対象者にとって明らかに容易すぎたり、困難すぎる項目についての評価を省略することができ、大幅な面接時間の短縮を実現している。この工夫によって、VABS-II は、ウェクスラー式の知能検査のような年代別のフォームではなく、全ての年代に共通のフォームで評価を行うことを可能にしている。

しかし、このルールの適用には、項目が困難度順に配置されていることが不可欠な条件となる。そうでなければ、評価を行っていない項目に「0」や「2」のスコアを割り当てることの正当性が保証されないためである。原版は、アメリカにおける調査で測定された困難度をもとに項目を配置しているが、文化的な差異を考えれば、日本における各項目の困難度がアメリカと同一であるとは考えにくいいため、日本での実証的な検討が必要である。

そこで本研究では、調査の実施段階では打ち切りルールを適用せず、全ての項目に回答を求めることによって、各項目についてより正確な困難度の評価を行うこととする。この結果に基づいて項目順の最終的な配置を決定することにより、VABS-II 日本

版の実施において打ち切りルールを適用するための前提条件を整えることができる。

## 7. その他の調査

尺度の基準関連妥当性を検証するため、VABS-IIの他に4つの調査を実施した。

### 1) 旭出式社会適応スキル検査

旭出式社会適応スキル検査（宇佐美他, 2011）は、「日常生活において機能するために人々が学習した、概念的、社会的および実用的なスキル」として定義される社会適応スキルを測定するために国内で開発された192項目の質問紙尺度である。内的整合性および再検査信頼性の観点から信頼性が確認されており、妥当性についても、年齢による得点推移や新版S-M社会生活能力検査との関連によって検証されている。

当検査は、言語スキル（基本的な言語の理解や表出に関わるスキルや読み書きのスキル）、日常生活スキル（家の中での生活に必要なスキル）、社会生活スキル（家の外や地域での生活に必要なスキル）、対人関係スキル（対人的なやりとりや集団参加に必要なスキル）の4領域から構成される。各項目について、「できる、今は行わないが小さいときはできていた（2点）」か「できない、やったことがない（0点）」のいずれかで評定するが、項目によっては、その中間として「1点」（状況によってできる、部分的にできるなど）が設けられている。得点が高いほど、適応スキルが高いことを意味する。

本研究では、標準化サンプルとは別の一般群の対象者について、保護者、配偶者、きょうだいなど対象者の日常をよく知る人物に当検査を実施した。当検査の適用年齢

は3-18歳とされているが、より幅広い年代での基準関連妥当性を検証するため、本研究では19歳以上の対象についても、当検査を実施した。

### 2) CBCL

Child Behavior Checklist (CBCL; Achenback, 1991)は、幼児期から青年期の情緒的・行動的問題を包括的に評価する118項目の質問紙尺度である。日本語版は井潤他(2001)によって開発され、国内で信頼性・妥当性が確認されている。

CBCLは、ひきこもり、身体的訴え、不安／抑うつ、社会性の問題、思考の問題、注意の問題、非行的行動、攻撃的行動の8つの下位尺度からなり、さらにひきこもり、身体的訴え、不安／抑うつからなる内向尺度と、非行的行動と攻撃的行動からなる外向尺度、全下位尺度の総得点から、総合的に適応上の問題を評価する。各項目について、「あてはまらない（0点）」、「ややまたはときどきあてはまる（1点）」、「よくあてはまる（2点）」の3段階で評定する。得点が高いほど、適応上の問題行動の頻度・程度が高いことを意味する。

本研究では、標準化サンプルとは別の一般群の対象者について、保護者、配偶者、きょうだいなど対象者の日常をよく知る人物に当検査を実施した。当検査の適用年齢は4-18歳とされているが、より幅広い年代での基準関連妥当性を検証するため、本研究では19歳以上の対象についても、当検査を実施した。

### 3) 知能指数 (IQ)

臨床群の対象者については、回答者への

聞き取りによって、ウェクスラー式またはビネー式知能検査による IQ の情報を得た。上記知能検査を受けた経験がない、または、正確な IQ 値について回答者の記憶がないもしくは曖昧なケースを除き、184 名 (71.0%) について IQ の情報が得られた ( $M = 74.2, SD = 32.0$ )。

#### 4) PARS

PARS (安達ら, 2006; 神尾ら, 2006; 辻井ら, 2006) は、国内で開発・標準化された ASD アセスメントのための半構造化面接形式の尺度であり、ASD の識別力や ASD アセスメントのゴールドスタンダードである ADI-R との関連などの観点から妥当性が確認されている (Ito et al., in press)。

PARS の実施に際しては、幼児期の最も症状の顕著だったときに関する回顧評定 (ピーク評定) と現在の症状に関する現在評定を求める。ピーク評定尺度は対象者の年齢を問わず同一の形式 (34 項目) をとるが、現在評定尺度は、幼児期 (34 項目)、児童期 (33 項目) および思春期・成人期 (33 項目) の各年齢帯に対応した 3 つのバージョンが存在する。また、ピーク評定尺度については因子構造が検討されており、社会性 (8 項目)、敏感性 (10 項目)、常同行動 (8 項目)、こだわり (8 項目) の 4 下位尺度が見出されている。

PARS の項目は、ASD 児者に特徴的な行動特徴を記述したもので、「なし (そのようなことはなかった/ない) (0 点)」、「多少目立つ (多少そのようなことがあった/ある) (1 点)」、「目立つ (よくそのようなことがあった/ある) (2 点)」の 3 段階で評定を行う。得点が高いほど、自閉的特性が

顕著であることを意味する。

本研究では、臨床群の対象者について、保護者、配偶者、きょうだい、施設職員等、対象者の日常の様子をよく知る者に対して PARS を実施した。

#### 8. 倫理面への配慮

本研究に関して、浜松医科大学の倫理委員会等における検討と承諾を得た。また、対象者については、調査の趣旨や個人情報の保護について十分に説明し、インフォームドコンセントを得た上で面接を行い、面接後には謝礼として図書カードを提供した。

#### 9. 分析

まず、IRT によって推定された困難度パラメータに基づき項目順の並び替えを行った上で、打ち切りルール (前述) を適用した。この打ち切り適用後の得点に基づいて、94 の年齢区分 (Table 3 参照) ごとに各下位領域の素点を標準得点に換算する数表を作成した。次に、VABS-II の信頼性について、再検査信頼性、評定者間信頼性の 2 つの観点から検討した。最後に、VABS-II の妥当性について、基準関連妥当性、臨床群のスコアプロファイルという 2 つの観点から検討を行った。なお、項目分析や内的整合性など、昨年度と重複する分析結果については、総合研究報告書にまとめて報告した。

### C. 研究結果

#### 1. VABS-II 日本版の標準化

##### 1) 標準化サンプル

Table 3 に尺度の標準化 (各尺度の素点を個人の年齢段階における相対的位置を示す標準得点に換算できるようにすること) に

用いたサンプルの内訳を示す。94の年齢区分ごとに信頼できる標準値が得られるよう、各年齢区分および性別の偏りができるだけ少なくなるようサンプルを収集した。

## 2) テスト情報量

IRT（項目反応理論）では情報量による測定精度の評価が得られる。古典的テスト理論における信頼性係数が尺度の測定精度を1つの値で評価し、受験者の特性値の分布（分散の大きさ）に影響されるのに対して、IRTにおける情報量は受験者の特性値の分布と独立に、特性レベルごとに測定精度が評価される。

Figure 1～Figure 11に、下位領域ごとのテスト情報量と、年齢帯ごとの特性値の分布（箱ひげ図）を示した。同時に、情報量のグラフの上部に、各年齢帯の特性値の中央値の位置を示した。特性値の分布と情報量を比較することによって、各尺度が測定精度を発揮する年齢帯を明らかにすることができる。下位領域によってどの範囲の年齢帯で高い測定精度が得られるかは異なっていることが分かる。受容言語、表出言語、微細運動、粗大運動などでは情報量が高いのは4, 5歳程度までであるのに対して、読み書き、地域生活、対人関係などはより高い年齢帯まで情報量が高い。前者の下位領域が早い年齢帯で発達する行動に敏感な尺度となっているのに対して、後者の下位領域は比較的高い年齢帯で発達する行動を評価する尺度となっていることが示される。また、すべての下位領域において、ある程度の年齢以上では測定精度は著しく低くなる。これは、どの領域においても年齢が進むとほぼすべての項目を通過（評価2が与

えられる）する者が多数を占め、天井効果が発生しているためである。特性値の箱ひげ図において、高い年齢帯で分布が同一の上限値に貼り付くようになるのも、そのためである。

## 3) 打ち切りルールに関する検討

本研究では、すべての対象者に全項目を実施し、その結果に基づいて各項目の困難度を決定した。しかしながら、実際にVABS-II日本語版を実施する際には、評価「0」が4項目続くとそれ以降の項目の評価は行わずにすべて「0」とし、また評価「2」が4項目続いた場合にはそれ以前の項目の評価を行わずすべて「2」とするという、打ち切りルールが用いられる。ここでは、前節で述べた困難度に基づく新しい項目配置において打ち切りルールを適用した場合の尺度得点と、打ち切りルールを適用せずに得られた尺度得点を比較することによって、打ち切りルール使用の妥当性を検討する。

Figure 12～Figure 15に、打ち切りルールを適用していない尺度得点と、打ち切りルールを適用した場合の尺度得点との散布図を示した。両者の相関は非常に高く（相関係数はすべて.99以上）、打ち切りによる測定精度の低下はほとんどないことが示された。このことから、困難度に基づく項目配置における打ち切りルール適用の妥当性が示された。

## 4) 標準得点の算出

分担研究報告書「V-scale 算出のための新たな試み」に示した系列範疇法に基づくアルゴリズムを用いて、前節で算出された打ち切り適用後の各下位領域得点を94の年

年齢区分 (Table 3 参照) ごとの標準得点 (個人の年齢段階における相対的位置を示す得点) に変換した。

この下位領域ごとの標準得点は、原版では V-scale と呼ばれ、15 を平均、3 を 1SD とする得点である。一般に、標準得点として Z 得点 (各個人の素点から平均値を減算し、標準偏差で除した値) が用いられるが、VABS-II のように全年代にわたる細かい年齢区分ごとに標準得点を算出する場合、各年齢区分のサンプルサイズが限られるため、年齢区分ごとの平均値や標準偏差の値が不安定になりやすい。また、VABS-II では特に低年齢帯や高年齢帯において、フロア効果や天井効果によって正規分布とはかけ離れた得点分布となるため、サンプルから得られた平均値や標準偏差から直接標準得点を算出することは望ましくない。そこで本研究では、系列範疇法に基づくアルゴリズムによって素点の累積度数分布を正規分布に近似した上で、各年齢区分の標準得点を算出する方法を用いた。詳細は分担研究報告書「V-scale 算出のための新たな試み」を参照されたい。

## 5) 標準得点の分布

Figure 16~Figure 34 に各下位領域の標準得点 (V-scale) の 4 年齢帯ごとの分布を示す。VABS-II の各下位領域は、一般的に青年期から成人期には素点の分布自体が右側に大きく偏ったものとなり、平均値も尺度の上限に近い値をとるようになるため (結果 2-2) に後述)、一定の年齢以上では素点が尺度の上限値であっても、標準得点としては平均を示す 15 点付近の得点を与えられる下位領域が多くなる。言い換えれば、多

くの下位領域では高年齢帯において理論上の標準得点の上限値が 15 点またはその付近になる。その結果、7-15 歳や 16-92 歳における標準得点の分布では、正規分布に比して 17-18 点以上の割合が小さくなり、その分、15-16 点の割合が大きくなっている下位領域が多い。特に受容言語、表出言語はその傾向が顕著である。

しかし、VABS-II はもともと健常者における適応行動の個人差を測定するためのツールではなく、様々な障害や疾患を抱える者の適応を評価するためのツールであるため、平均以上の得点を示す人々の間での個人差を捉えられないことによって直接的な支障が生じるわけではない。したがって、15 点以上よりも、むしろ 15 点未満の得点帯で分布の正規性が保たれているかに着目することが重要である。その視点に立って 15 点未満の分布を見ると、いずれの下位領域・年齢帯も平均-2SD にあたる 9 点以上の範囲に大部分のケースが含まれており、分布の形状もおおむね滑らかで正規分布に近い様相を呈している。このことから、15 点未満の標準得点に関しては、下位領域や年齢を超えた共通の尺度として利用することの妥当性が示されたと言える。以下の分析では、特に記載がない場合、この標準得点を使用して分析を行った。

## 2. VABS-II 日本版の信頼性

### 1) 再検査信頼性

再検査信頼性の検証に際しては、標準化サンプルとは別に男女 45 名のデータを得た (Table 18)。同一の対象者について、同一の評定者が 2 週間から 1 ヶ月の期間を空けて 2 回の面接調査を実施した。

下位領域・領域ごとの標準得点における2回の調査の相関を Table 19 に示す。いずれの下位領域・領域も 0.1%水準で有意な相関を示している。係数は一部に.70を下回る下位領域が見られるが、大部分は.70～.90の範囲にあり、十分な再検査信頼性が示されたと言える。

## 2) 評定者間信頼性

再検査信頼性と同様、標準化サンプルとは別に14名のデータを得た (Table 18)。同一の対象者について、異なる評定者が1ヶ月以内に2回の面接調査を実施した。

下位領域・領域ごとの標準得点における2回の調査の相関を Table 19 に示す。不適応外在化を除く全ての下位領域で有意な相関が示された。係数は全般的に再検査信頼性とほぼ同等の値を示しており、十分な評定者間信頼性が示されたと言える。ただし、身辺自立や不適応外在化は低い係数を示している。この結果は、サンプルサイズの不足による相関の不安定さに起因する可能性もあり、今後、データを追加して、より安定的な結果を得る必要がある。

## 3. VABS-II 日本版の妥当性

### 1) 旭出式検査・CBCL との関連

VABS-II の基準関連妥当性を検討するため、標準化サンプルとは別の一般群を対象として旭出式社会適応スキル検査およびCBCLを実施した。サンプルの内訳を Table 25 に示す。

旭出式検査は適応行動と概念的な関連の深い社会適応スキルを測定するために開発された尺度である。したがって、VABS-II の適応行動尺度と強い相関を示す一方、不

適応行動尺度との関連は弱いものに留まることが予測される。また、旭出式検査は言語スキル、日常生活スキル、社会生活スキル、対人関係スキルの4領域に分かれている。概念的な関連から、言語スキルはVABS-II のコミュニケーション領域、日常生活スキルはVABS-II の日常生活スキル領域、社会生活スキルおよび対人関係スキルはVABS-II の社会性領域と相対的に強い関連を示すと考えられる。ただし、「因子的妥当性」の項で示したように適応行動の各領域は相互に強い関連を有するため、上に示した組み合わせ以外でも一定の相関が見られることが予測される。

VABS-II と旭出式検査はともに年齢の影響を強く受けるため、年齢を統制変数とした素得点間の偏相関を算出した。ただし、旭出式検査の対象年齢は18歳までであり、19歳以降は顕著な得点の上昇がないと考えられるため、19歳以降の対象者については年齢を18歳と見なして分析を行った。

Table 26 にVABS-II の各領域・下位領域と旭出式検査の各領域間の偏相関を示す。旭出式検査の各領域は全般的に適応行動尺度の各下位領域および領域とは強い正の相関を示しているが、不適応行動尺度とは有意な相関を示していない。また、領域ごとの差異に着目すると、旭出式検査の言語スキルはVABS-II のコミュニケーション領域、日常生活スキル、社会生活スキルおよび対人関係スキルの3領域はVABS-II の社会性領域と最も強い相関を示している。この結果から、概念的関連に基づく予測はおおむね支持され、VABS-II の収束的・弁別的妥当性が示された。ただし、日常生活スキルに関しては、旭出式検査の日常生活スキル

と相対的に弱い関連を示している。これはサンプルサイズの不足による相関の不安定さに起因する可能性もあるため、今後、データ収集を継続して、より安定的な結果を得る必要がある。

CBCL は VABS-II の不適応行動と関連の深い情緒的・行動的問題を把握するために開発された尺度である。したがって、VABS-II の不適応行動尺と強い関連を示す一方、適応行動との関連は限定的であることが予測される。また、CBCL は 8 つの下位尺度からなり、そのうち、ひきこもり、身体的訴え、不安／抑うつが内向尺度、非行的行動と攻撃的行動が外向尺度を構成している。これらはそれぞれ VABS-II の不適応内在化および不適応外在化と相対的に強い相関を示すことが予測される。

Table 27 に VABS-II の各領域・下位領域と CBCL の各下位尺度の相関を示す。全体的に、CBCL の各下位尺度は、VABS-II の適応行動とは一部弱い相関を示すのみに留まっているが、不適応行動とは強い相関を示している。尺度間の差異を見ると、内向尺度に属するひきこもり、身体的訴え、不安／抑うつや内向尺度総得点は VABS-II の不適応内在化と相対的に強い相関を示している。一方、外向尺度に属する非行的行動、攻撃的行動や外向尺度総得点は不適応外在化と相対的に強い相関を示している。これらの結果から、概念的関連に基づく予測はいずれも支持され、VABS-II の収束的・弁別的妥当性が示された。

## 2) 臨床群のスコアプロフィール

VABS-II によって発達障害・知的障害を含む様々な障害を抱える者の適応上の困難

を把握しうるかについて検討するため、臨床群を対象に VABS-II を実施した。臨床群サンプルの内訳は、MR 群（知的障害あり、発達障害なし）が 38 名、HFASD 群（知的障害なし、ASD）が 80 名、LFASD 群（知的障害あり、ASD）が 84 名、ADHD 群（知的障害なし、ADHD）が 12 名、視聴覚障害群（視聴覚障害のみ、知的障害・発達障害なし）が 8 名、その他（LD、知的水準不明の ASD や ADHD など）が 37 名であった（Table 28）。各群の平均年齢および IQ を Table 29 に示す。いずれの群も平均年齢は 10-15 歳程度であった。平均 IQ は MR 群と LFASD 群が約 50、HFASD 群と ADHD 群が 100-110 程度であった。視聴覚障害群およびその他については IQ の情報が得られなかったが、知的障害の診断を受けているケースは視聴覚障害群から除外している。

DSM-IV（APA, 2000）に記されているように知的障害は特定の領域ではなく適応行動全般に影響を及ぼす障害であるため、MR 群は適応行動尺度の全領域の得点が同等に低いプロフィールを示すと考えられる。一方、発達障害を合併していない場合、不適応的な問題行動の頻度・程度は低いため、不適応行動尺度の得点は ASD 群や ADHD 群よりも低いと予測される。

ASD は、社会性やコミュニケーションの困難を主症状とする障害であるため、HFASD 群や LFASD 群は特に社会性領域やコミュニケーション領域の得点が相対的に低いプロフィールを示すと考えられる。ただし、HFASD 群については、言語に障害のないアスペルガー症候群のケースが多く含まれるため、コミュニケーション領域における得点の低さは見られない可能性もある。

不適応行動については、特に ASD の行動症状と関連の深い内在化問題において MR 群よりも高い得点を示すと考えられる。

ADHD は不注意や多動性を主症状とする障害であるが、ASD との関連の深さも指摘されているため、ADHD 群は社会性領域で他領域よりもやや低い得点を示すと予想される。また、不適応行動については、外在化問題において MR 群や ASD 群より高い得点を示すと考えられる。

視聴覚障害は生活全般への適応を困難にすると考えられるため、全般的に得点の低いプロフィールを示すと考えられる。一方、不適応行動については、他のいずれの群よりも低い得点を示すと考えられる。

Table 30 に各群の VABS-II および PARS の尺度得点の記述統計を、Figure 60、Figure 61、Figure 62 に各群の VABS-II のスコアプロフィールを示す。また、Figure 63～81 に各群における VABS-II の尺度得点の分布を示す。適応行動尺度の領域ごとのスコアプロフィールを見ると、全体的な水準としては、HFASD 群および ADHD 群が比較的高く、MR 群、LFASD 群、視聴覚障害群は低水準にある。しかし、HFASD 群と ADHD 群も一般群の平均値である 15 点よりほどの領域も 1SD (3 点) 程度低く、特に HFASD 群では、社会性の落ち込みが顕著で、平均 -2SD にあたる約 9 点を示している。一方、MR 群は全般的に平均 -3SD 程度の得点を示している。同程度の知的水準にある LFASD 群では全般的な水準は MR 群と同等であるが、コミュニケーションと社会性が MR 群より 1～2 点ほど低い値を示しており、ASD の特徴が表れている。視聴覚障害群は、全般的に水準が低いが、LFASD 群とは対照的

に、コミュニケーションや社会性の得点が相対的に高いプロフィールを示している。

一方、不適応行動尺度の得点を見ると、MR 群や視聴覚障害群に比べて、3 つの発達障害群が高い得点を示している。HFASD 群と LFASD 群はいずれも内在化得点が外在化得点より高いが、LFASD 群の方がいずれも高い水準を示している。それに対し、ADHD 群は外在化得点が内在化得点より高いという対照的なプロフィールを示している。適応行動尺度では HFASD 群と ADHD 群の差異はそれほど明確でなかったが、不適応行動尺度ではこのような顕著な差異が表れている。

以上の結果から、VABS-II の適応行動尺度および不適応行動尺度を組み合わせることで、発達障害を含む様々な障害に起因する適応上の問題の実態を把握可能であることが明らかになった。各群のスコアプロフィールは各障害の概念的特徴を明確に反映しており、VABS-II の高い妥当性が示された。

### 3) 知能指数・自閉症特性との関連

前節では障害群ごとの差異を検討したが、同一の診断カテゴリに属するケースの中でも知的水準や症状の表れ方には大きな個人差が存在する。こうした個人差と適応行動・不適応行動の連続的な関連性を検討するため、臨床群全体における知能指数および自閉症特性と VABS-II の相関を検討した。

前節でも示されたように知的障害は全般的な適応を阻害するため、IQ は適応行動の全ての領域と正の関連を示すと予測される。一方、知的障害が直接不適応的な問題行動につながるとは限らないため、不適応行動



との相関は弱いものに留まると考えられる。

自閉症特性は、コミュニケーションや社会性の障害をもたらすため、適応行動尺度のコミュニケーション領域や社会性領域の得点と負の関連を示すと予想される。また、不適応行動尺度とは、内在化問題を中心に正の関連が見られると予測される。

Table 31 に VABS-II の各領域・下位領域得点と IQ および PARS の各尺度得点の相関を示す。IQ との相関を見ると、適応行動尺度の全ての領域・下位領域が IQ と中程度～強い正の相関を示した。領域別では、コミュニケーションが最も係数が高く、日常生活、社会性、運動スキルと続く。総得点との相関は.72 で、IQ によって総得点の約 50% の分散が説明されることが示された。一方、不適応行動については内在化問題と総得点のみ弱い負の相関を示し、外在化問題は有意な相関を示さなかった。

PARS と適応行動尺度の相関については、PARS ピーク評定の 4 因子のうち社会性因子や常同行動因子が相対的に強い相関を示した。特に VABS-II のコミュニケーション領域や社会性領域との関連が強かった。一方、PARS 現在評定は適応行動尺度との相関が弱く、青年成人期尺度と社会性領域の間に弱い相関が見られるに留まった。

PARS と不適応行動尺度の相関を見ると、PARS ピーク評定は内在化問題と弱い～中程度の相関を示した。適応行動とは対照的に、4 因子のうち過敏性因子やこだわり因子が相対的に高い係数を示した。また、現在評定も児童期、青年成人期ともに内在化問題と中程度の相関を示し、青年成人期では外在化問題との関連も見られた。全ての下位尺度が適応行動もしくは不適応行動の

どちらかと関連を示したことは興味深い。この結果は、自閉症特性の中でも適応行動の減少につながる側面と、不適応行動の増加につながる側面が存在することを示唆している。

以上をまとめると、IQ は VABS-II の適応行動と強い相関、不適応行動の内在化問題と弱い相関を持つこと、自閉症特性は発達初期の社会性や常同行動が適応行動と中程度～強い相関、発達初期の過敏性、こだわりや現在の行動症状が不適応行動の内在化問題と中程度の相関を持つことが示された。これらの相関のパターンは概念的な予測とおおむね一致しており、VABS-II の基準関連妥当性の高さが示された。

## D. 考察

### 1. VABS-II 日本版の標準化に関して

IRT によって推定されたテスト情報量の分布と年齢帯ごとの特性値から、適応行動尺度の各下位領域は、互いに異なる年齢帯の対象に対して最大のテスト情報量を発揮することが示された。例えば、下位領域受容言語では 1-3 歳付近でテスト情報量がピークとなるが、読み書きでは 4-15 歳程度までが情報量のピークとなっている。このように VABS-II は、互いに異なる年代の個人差に敏感な複数の下位尺度を組み合わせることで、乳幼児から成人に至るまでの幅広い年代における適応行動の発達を評価することができるように構成されていることがわかる。

VABS-II の適応行動尺度では、実施時間の短縮のため、各下位領域において 4 つ以上連続して「0」や「2」の評定が続いた場合には、評定を打ち切って、それ以降（ま

たはそれ以前)の項目を全て「0」または「2」と見なすという打ち切りルールが存在する。このルールが有効に機能するためには、各下位領域内で項目が困難度順に配置されていることが前提条件となる。そこで本研究では各項目の困難度を正確に推定するため、調査の実施段階では打ち切りルールを適用せず、全項目について聴取を実施した。こうして測定された全項目の評定データに基づいて、IRTによって各項目の困難度を推定した結果、全体的には原版の項目順と調査データの困難度順は一致していたが、一部に困難度の逆転している箇所が見られたため、困難度に基づく項目順の再配置を行った(新旧の項目配置についてはTable 3を参照)。この再配置後の項目順において打ち切りルールを適用し、打ち切り前後の得点の相関を検討したところ、いずれの下位領域についても.99程度のきわめて強い相関が見られた。このことから、新しい項目順において、打ち切りによる誤差の混入は最小限に留まることが明らかになり、打ち切りルール適用の妥当性が示された。これによって、VABS-II 日本版においても打ち切りルールを使用して原版と同様の20-60分程度の時間で聴取を実施することが可能となった。

この打ち切り適用後の得点を用いて、尺度の標準化(各尺度の素点を個人の年齢段階における相対的位置を示す標準得点に換算できるようにすること)を行った。知能検査におけるIQと同様、VABS-IIにおいては94の年齢区分ごとの標準値に基づいて標準得点が算出される。この標準値を得るために本研究では0-92歳までの全年代にわたる計1367名のデータを収集した。しかし、

各年齢帯における素点の分布は、特に低年齢帯や高年齢帯では、フロア効果や天井効果の影響で正規分布をなさないため、一般的なZ得点(偏差値)の算出のように、素点の平均値と標準偏差から直接標準得点を算出することができない。そのため、何らかの形で各年齢帯の素点分布を非線形的に正規分布に変換するアルゴリズムが必要となる。原版ではこの問題を解決するために、Johnson 曲線適合プログラム(Hill, Hill & Holder, 1976)を用いているが、このプログラムは素点の平均値、標準偏差、歪度、尖度という4つの積率に基づく簡便なアルゴリズムであり、一般に3次積率(歪度)や4次積率(尖度)がきわめて不安定な性質を持つ(サンプルの偏りによる影響を受けやすい)ことを考慮すると、推定結果には相当の誤差が混入することが予想され、最善の方法とは考えにくい。そこで、本研究では最小二乗法による系列範疇法(Diederichi, Messick, & Tucker, 1957)に基づいて、素点の累積度数分布を正規分布に近似するアルゴリズムを適用した。

このアルゴリズムに基づいて算出された標準得点(V-scale)の分布を確認したところ、理論上の標準得点の上限値が制限される高年齢帯を除いて、ほぼ正規分布に近い分布が得られた。高年齢帯でも、平均にあたる15点を下回る得点帯ではおおむね正規分布に近い分布が示された。もともとVABS-IIは障害や疾患による適応上の困難を把握するための尺度であるため、平均以上の得点帯における個人差を敏感に検出する必要性は低い。平均未満の得点帯で分布の正規性が確認されたことで、少なくとも年齢段階の標準よりも低い適応行動を示す

ケースにおいては、V-scale が年齢や尺度を超えた共通の評価基準として機能することが示された。

## 2. VABS-II 日本版の信頼性に関して

再検査信頼性および評定者間信頼性について検討を行った結果、一部の低位領域でやや低い係数が見られた他は、十分な信頼性が確認された。このことから、VABS-II の実施において生じる同一評定者内でのランダムな測定誤差や評定者間での測定値のズレは最小限であることが明らかになった。ただし、評定者間信頼性については、サンプルサイズが不足していることもあってか、低位領域によるバラつきがやや大きく、今後、データを追加して、より安定的な結果を得る必要がある。

## 3. VABS-II 日本版の妥当性に関して

VABS-II は適応に必要なスキルの発達を評価する適応行動尺度と適応を阻害する問題行動の出現を評価する不適応行動尺度から構成される。これら2つの尺度の基準関連妥当性を検証するため、旭出式社会適応スキル検査およびCBCLとの関連を検討した。旭出式検査は適応行動と概念的な関連の深い社会適応スキルを測定するための検査であり、VABS-II の適応行動尺度との強い関連が予測される。一方、CBCL は適応上の情緒的・行動的問題を包括的に測定する尺度であり、VABS-II の不適応行動尺度との関連が予想される。分析の結果、予測に一致し、旭出式検査は適応行動尺度と、CBCL は不適応行動尺度と強い相関を示した。また、低位尺度単位でも概念的に関連の深い低位尺度同士が相対的に強い相

関を示し、VABS-II 日本版の各領域・低位領域の収束的・弁別的妥当性が示された。

VABS-II は様々な障害や疾患による適応上の問題を把握するために開発された尺度である。そこで、知的障害(MR)、ASD、ADHD、視聴覚障害という4つの障害による適応上の問題をVABS-II 日本版によって把握しうるか否かをスコアプロフィールから検討した。その結果、MR 群は適応行動全般の得点が低い一方、不適応行動では発達障害群ほど高い得点を示さないことが明らかになった。一方、ASD 群は知的水準によって予測されるよりも低い適応行動を示し、特に社会性領域で低い得点が示された。また、不適応行動では内在化問題が顕著であった。ADHD 群は、適応行動に関してはASD 群と類似したパターンを示したが、不適応行動では、ASD 群と対照的に外在化問題が顕著な高得点を示した。視聴覚障害群は、全般的にMR 群と同等の低い適応行動得点を示したが、ASD 群とは対照的にコミュニケーションや社会性領域で相対的に高い得点を示した。不適応行動は他のいずれの群よりも低い得点を示した。これらの結果から、VABS-II 日本版の適応行動尺度と不適応行動尺度を組み合わせることで、様々な障害の特徴を明確に把握できることが明らかになった。

知的障害や発達障害は症状の個人差がきわめて大きく、同一の診断カテゴリに属していても、その適応の程度は個人によって様々である。そこで、臨床群全体における知的水準および自閉症特性の量的程度とVABS-II スコアの連続的な関連について検討した。その結果、上述の障害群ごとの分析結果と一致して、IQ は適応行動全般の得

点と正の相関を示し、不適応行動とは内在化問題と弱い負の相関を示した。また、PARSによって測定された自閉症特性は、幼児期ピーク評定の社会性因子や常同行動因子が適応行動尺度全般と負の相関、過敏性因子、こだわり因子や現在評定が不適応行動の内在化問題と正の相関を示した。これらの結果から、VABS-IIの適応行動尺度と不適応行動尺度は、IQや自閉症特性の量的程度を明確に反映することが明らかになり、VABS-IIを用いることで診断カテゴリに依存しない、個々のケースごとの量的な適応評価を実現可能であることが示唆された。

#### 4. 総括と今後の展望

発達障害児者への具体的な支援計画の策定や行政サービスの実施を考える上で、個々の支援ニーズを的確に把握することが必要となるが、国内では発達障害児者の支援ニーズを客観的に把握するための基本的なツールが開発・普及されておらず、もっぱら知的機能に基づいた評価のみが行われてきた現状がある。しかし、高機能ASD児者のように、知的機能には問題がなくとも、日常生活への適応に著しい困難を示すケースが少なからず存在しており、適応行動の観点から発達障害児者の支援ニーズを測定するツールの開発が急務とされている。そこで本研究では、適応行動の評価ツールとして国際的に最も広く利用されているVABS-IIの日本版を開発し、その標準化および信頼性・妥当性の検証を行った。

IRTに基づく項目分析と項目順の再配置を行った上で各下位領域の尺度得点を算出し、系列範疇法に基づくアルゴリズムによって標準得点を算出した。この標準得点は、

少なくとも障害児者のアセスメントにおいて重要となる平均未満の得点帯では分布の正規性を有することが確認され、年齢や尺度を超えた共通の評価基準として機能することが示された。内的整合性、再検査信頼性、評定者間信頼性の3つの観点から総合的に信頼性の検討を行った結果、VABS-II日本版が十分な信頼性を持つことが示された。また、妥当性について、因子的妥当性、年齢による推移、関連尺度との相関、臨床群ごとのスコアプロフィール、知的機能・自閉症特性との関連という5つの観点から、VABS-IIが高い妥当性を有し、発達障害を含む様々な障害を有する人々の個々の支援ニーズのアセスメントに効果を発揮することが示された。

今後の課題として以下の3点があげられる。1点目に、本研究では評定者間信頼性の検証や旭出式社会適応スキル検査との関連の検討において、十分なサンプルが得られなかった。今後、追加データを収集して、さらに安定的な結果を得る必要がある。2点目に、本研究では適応行動尺度の下位領域単位の標準得点(V-scale)を算出したが、領域単位の標準得点はV-scaleの平均値によって代用した。原版では領域全体のV-scaleの合計値を再度、領域標準得点に変換するアルゴリズムが用いられており、この点について今後検討を行う必要がある。また、不適応行動尺度についても標準得点の算出方法を検討する必要がある。3点目に、VABS-II日本版の出版化に向けて、発達障害児者に対する適切なアセスメントのためのガイドラインを作成していく必要がある。VABS-II日本版は、来年7月までの出版化の契約をすでに交わしており、今後の発達

障害児者の医療的・行政的支援において重要な役割を果たしていくと考えられる。また、この10年で発達障害の診断・評価に関する多くのアセスメントツールが開発された。今後、これらのツールと VABS-II を適切に使用し、個々の発達障害児者に対する具体的な個別支援計画を策定するためのガイドラインの作成が求められる。

#### E. 引用文献

- Achenbach, T. M. (1991) *Integrative Guide to the 1991 CBCL/4-18, YSR, and TRF Profiles*. Burlington, VT: University of Vermont, Department of Psychology
- 安達潤・行廣隆次・井上雅彦・他 (2006). 日本自閉症協会広汎性発達障害評価尺度 (PARS) ・児童期尺度の信頼性と妥当性の検討 臨床精神医学, 35, 1591-1599.
- American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., Text Revision) (DSM-IV-TR). Washington, D.C.: American Psychiatric Association.
- Diederich, G.W., Messick, S.J. & Tucker, L.R. (1957). A general least squares solution for successive intervals. *Psychometrika*, 22, 159-173.
- Hill, D., Hill, R., & Holder, R. L. (1976). Algorithm AS99: Fitting Johnson curves by moments. *Applied Statistics*, 45, 171-182.
- 井潤知美・上林靖子・中田洋二郎・北道子・藤井浩子・倉本英彦・根岸敬矩・手塚光喜・岡田愛香・名取宏美 (2001). Child Behavior Checklist/4-18 日本語版の開発 小児の精神と神経, 41, 243-252.
- Ito H, Tani I, Yukihiro R, Adachi J, et al. (In press). Validation of an Interview-Based Rating Scale Developed in Japan for Pervasive Developmental Disorders. *Research in Autism Spectrum Disorders*.
- 神尾陽子・行廣隆次・安達潤・他 (2006). 思春期から成人期における広汎性発達障害の行動チェックリスト：日本自閉症協会広汎性発達障害評価尺度 (PARS) の信頼性・妥当性についての検討 精神医学, 48, 495-505.
- Klin, A., Saulnier, C. A., Sparrow, S. S., Cicchetti, D. V., Volkmar, F. R., & Lord, C. (2007). Social and communication abilities and disabilities in higher functioning individuals with autism spectrum disorders: The Vineland and the ADOS. *Journal of Autism and Developmental Disorders*, 37, 748-759.
- Linacre, M. J. (2003). *WINSTEPS Rasch measurement software* (Computer software). Chicago, IL: Western Psychological Services.
- Linacre, J. M. (2012). *Winsteps® Rasch measurement computer program*. Beaverton, Oregon: Winsteps.com.
- Sparrow, S. S. Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service, Inc.
- Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales, Second Edition, Survey Forms Manual*. Minneapolis, MN: NCS Pearson, Inc.
- 辻井正次・行廣隆次・安達潤・他 (2006). 日本自閉症協会広汎性発達障害評価尺度

(PARS) 幼児期尺度の信頼性・妥当性の  
検討 臨床精神医学, **35**, 1119-1126.

Tomanik, S. S., Pearson, D. A., Loveland,  
K. A., Lane, D. M., & Shaw, J. B. (2007).  
Improving the reliability of autism  
diagnoses: Examining the utility of  
adaptive behavior. *Journal of Autism  
and Developmental Disorders*, **37**,  
921-928.

宇佐美慧・名越斉子・肥田野直・菊池けい  
子・斉藤佐和子・服部由紀子・松田祥子  
(2011). 社会適応スキル検査の作成の試  
み—検査の信頼性・妥当性・臨床的有用

性の検討— 教育心理学研究, **59**, 278-294.

F. 健康危険情報

特になし

G. 研究発表

総括研究報告書に記載

H. 知的財産権の出願・登録状況

特になし

Table 1. VABS-II 日本版における原版からの主な変更点

番号	変更前	変更後
表出言語		
32	現在進行形を使う	進行中の出来事や物事の状態を表したい時に「～ている」を使う(例えば、「雨がふっている」「おじさんが走っている」など)
34	代名詞を使う、性別にあった代名詞を使う必要があるが、文が文法的に正しい必要はない	代名詞を使う、文が文法的に正しい必要はない(例えば、「ぼくがやった」「これ、ください」などがある)
36	規則動詞の過去形を使う、不規則動詞は、文法的に誤って使うかもしれない	現在と過去を区別して表現する(例えば、「昨日、遊園地へ行った」「今日は幼稚園がある」など)
50	不規則の複数形を使う	助数詞をほぼ正しく使う(例えば、「1頭」「2羽」「車が3台」など)
読み書き		
1	1個のアルファベットを文字とわかり数字を区別する	1個以上の平仮名を文字とわかり数字と区別する
6	全ての活字体の大文字と小文字のアルファベットがわかる	すべての平仮名がわかる
9	追加	「がっこう」などの促音や「あそんだ」などの撥音が読める
10	少なくとも10の単語を声を出して読む	少なくとも10の漢字を覚えて書く
12	追加	50以上の漢字を覚えて書く
14	追加	「～は(wa)」「～へ(e)」といった特殊な読み方ができる
15	アルファベット順に単語を並べる	あいうえお順に単語を並べる
身辺自立		
18	正しくスプーン・フォーク・ナイフを使う	箸で食べる(こぼしてもよい、持ち方は問わない)
29	追加	こぼさないで箸が使える(握り箸ではない)

Table 2. VABS-II の尺度構成

行動	領域	下位領域	項目
適応行動	コミュニケーション	受容言語	20
		表出言語	54
		読み書き	25
	日常生活スキル	身辺自立	43
		家事	24
		地域生活	44
		社会性	38
	運動スキル	対人関係	31
		遊びと余暇	31
		コーピング	30
		粗大運動	40
		微細運動	36
不適応行動	不適応	内在化	11
		外在化	10
		その他	15
		重要事項	14

Table 3. 標準化サンプルの内訳 (次ページに続く)

年齢区分	男性	女性	合計	年齢区分	男性	女性	合計
0:0	8	8	16	2:0-2:1	13	13	26
0:1	11	15	26	2:2-2:3	5	5	10
0:2	6	9	15	2:4-2:5	5	7	12
0:3	8	7	15	2:6-2:7	5	4	9
0:4	10	9	19	2:8-2:9	6	4	10
0:5	7	10	17	2:10-2:11	5	5	10
0:6	12	13	25	3:0-3:1	5	7	12
0:7	9	8	17	3:2-3:3	5	6	11
0:8	7	6	13	3:4-3:5	6	5	11
0:9	12	7	19	3:6-3:7	5	5	10
0:10	8	7	15	3:8-3:9	8	7	15
0:11	11	14	25	3:10-3:11	2	5	7
1:0	11	9	20	4:0-4:1	5	6	11
1:1	10	11	21	4:2-4:3	5	10	15
1:2	12	15	27	4:4-4:5	7	8	15
1:3	7	7	14	4:6-4:7	6	9	15
1:4	11	9	20	4:8-4:9	5	8	13
1:5	9	6	15	4:10-4:11	4	3	7
1:6	11	9	20	5:0-5:1	5	4	9
1:7	12	10	22	5:2-5:3	5	5	10
1:8	9	11	20	5:4-5:5	5	3	8
1:9	13	7	20	5:6-5:7	7	4	11
1:10	9	6	15	5:8-5:9	4	4	8
1:11	5	5	10	5:10-5:11	6	4	10



Table 3. 標準化サンプルの内訳（続き）

年齢区分	男性	女性	合計	年齢区分	男性	女性	合計
6:0-6:2	7	8	15	13:0-13:5	6	7	13
6:3-6:5	8	7	15	13:6-13:11	10	9	19
6:6-6:8	8	1	9	14:0-14:5	4	3	7
6:9-6:11	8	4	12	14:6-14:11	6	8	14
7:0-7:2	3	6	9	15:0-15:5	6	4	10
7:3-7:5	5	6	11	15:6-15:11	3	8	11
7:6-7:8	3	5	8	16:0-16:5	5	6	11
7:9-7:11	7	7	14	16:6-16:11	8	5	13
8:0-8:2	4	5	9	17:0-17:5	4	4	8
8:3-8:5	8	4	12	17:6-17:11	4	5	9
8:6-8:8	7	8	15	18:0-18:5	6	4	10
8:9-8:11	4	5	9	18:6-18:11	3	4	7
9:0-9:3	4	7	11	19-21	11	17	28
9:4-9:7	11	2	13	22-29	27	22	49
9:8-9:11	11	12	23	30-39	15	13	28
10:0-10:3	6	3	9	40-49	7	13	20
10:4-10:7	6	5	11	50-54	6	10	16
10:8-10:11	5	7	12	55-59	9	7	16
11:0-11:3	8	4	12	60-64	3	6	9
11:4-11:7	7	6	13	65-69	5	3	8
11:8-11:11	9	8	17	70-79	10	9	19
12:0-12:3	8	4	12	80-92	8	14	22
12:4-12:7	8	8	16	合計	690	677	1367
12:8-12:11	7	5	12				

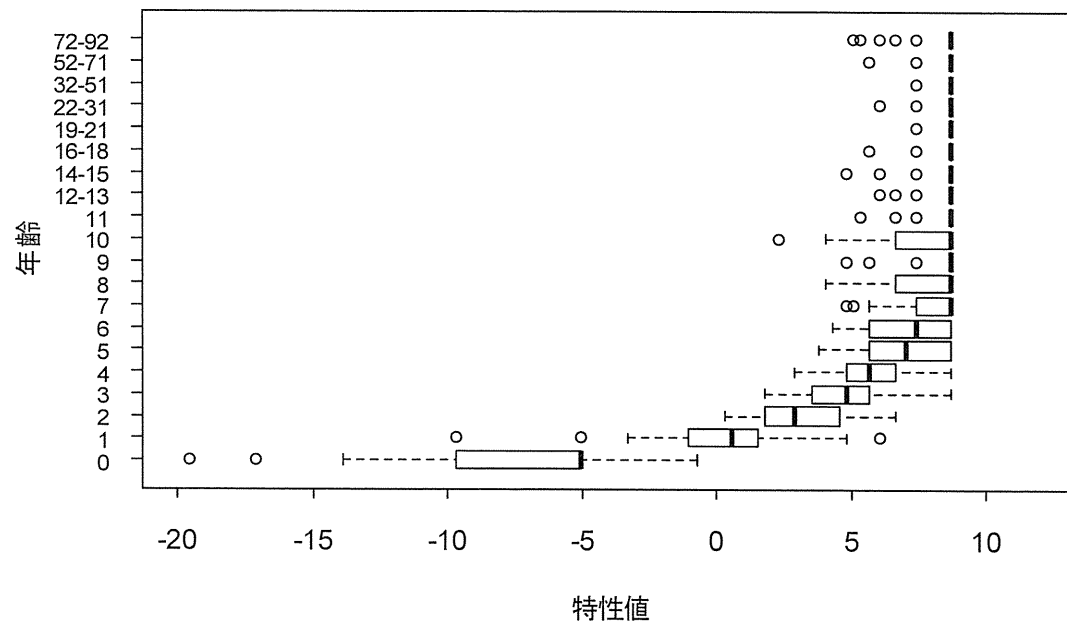
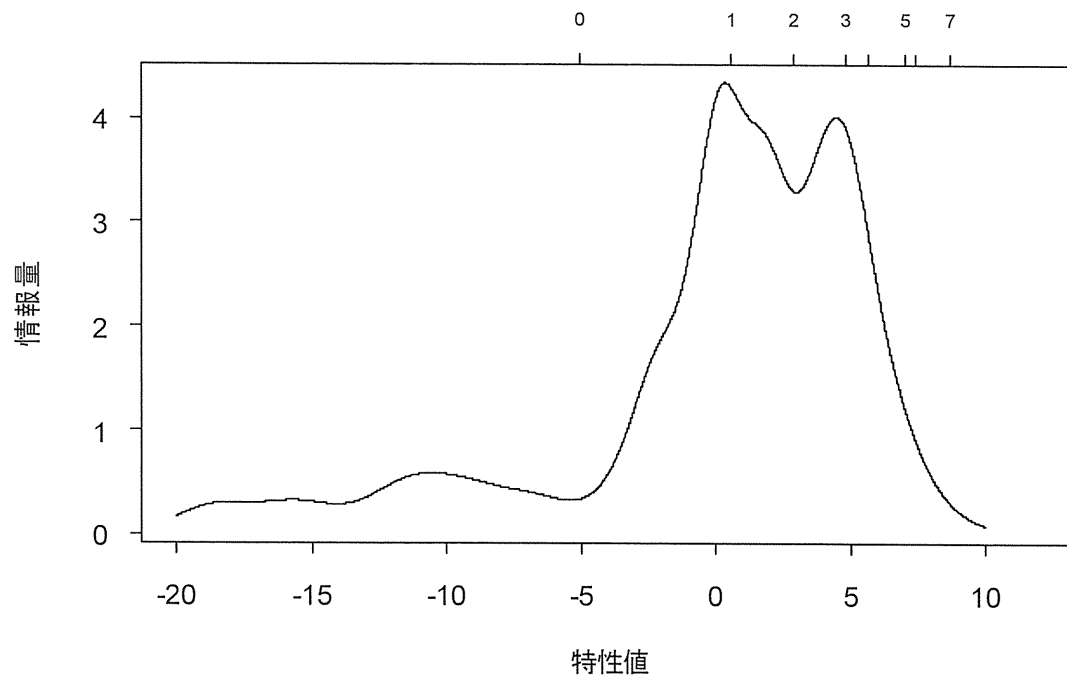


Figure 1. 受容言語のテスト情報量（上段）および年齢帯ごとの特性値の箱ひげ図（下段）

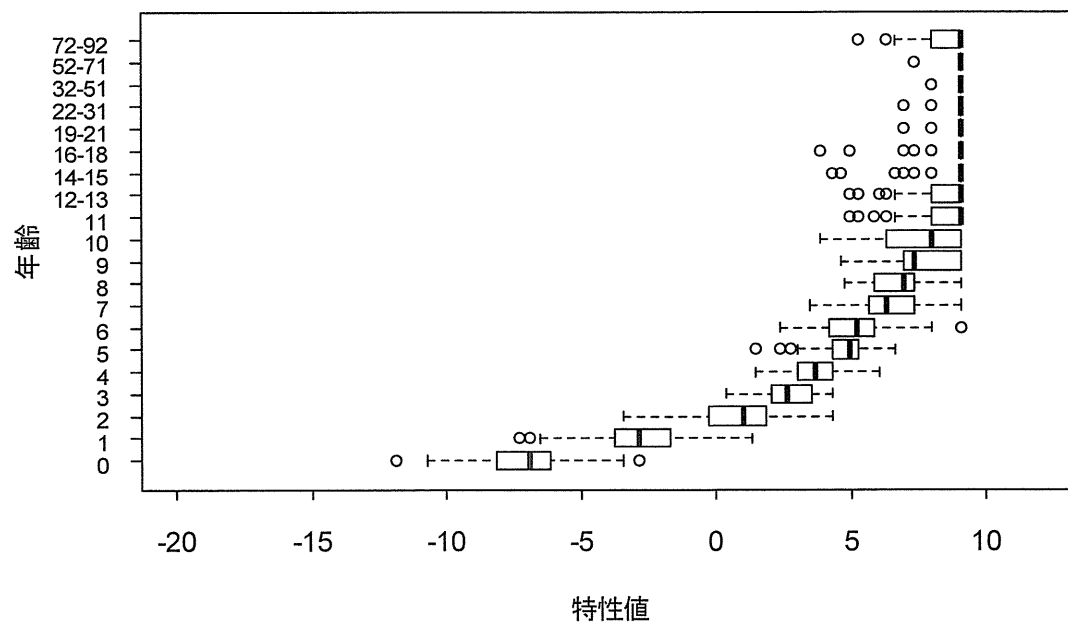
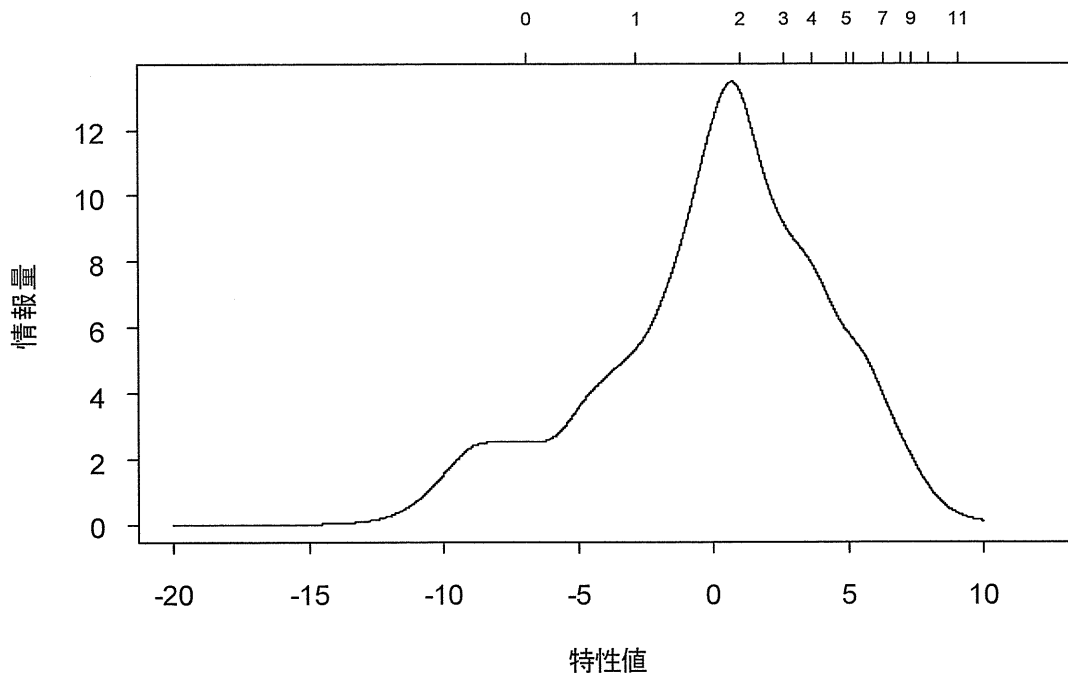


Figure 2. 表出言語のテスト情報量（上段）および年齢帯ごとの特性値の箱ひげ図（下段）

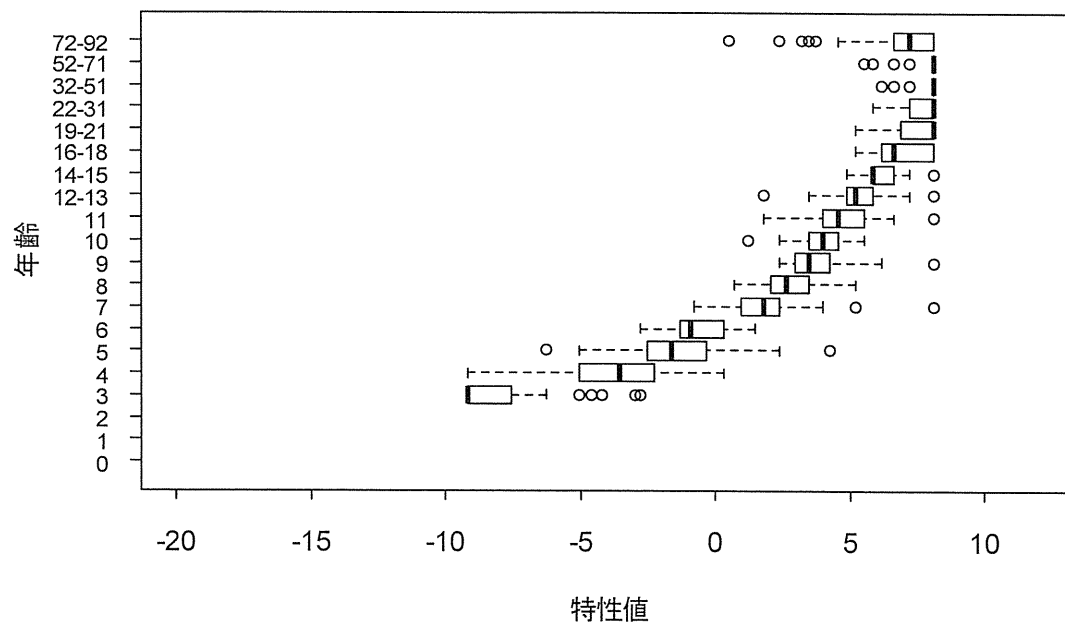
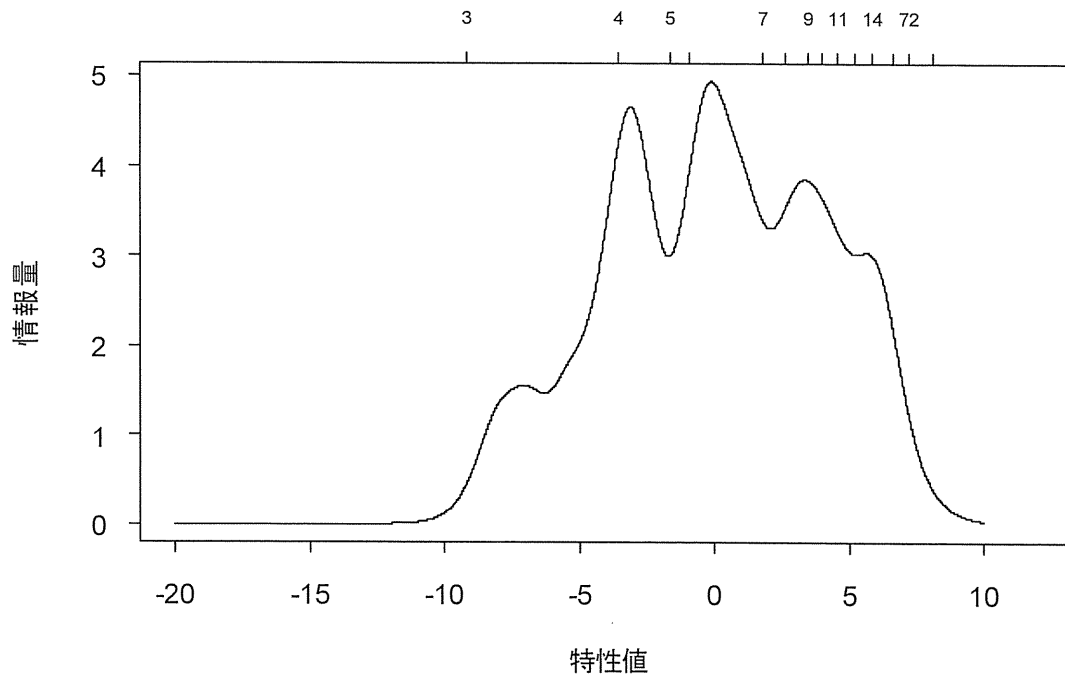


Figure 3. 読み書きのテスト情報量（上段）および年齢帯ごとの特性値の箱ひげ図（下段）