

内的整合性は、尺度を構成する複数の項目同士の関連の程度を意味する。再検査信頼性は、同一の検査を一定の期間を空けて二度実施したときの得点の相関の程度を指す。評定者間信頼性は、異なる評定者が同一の対象者に検査を実施したときの得点の相関の程度を意味する。これら3つの側面から検討を行うことで、項目間の測定対象のズレによる誤差、時間的な変動による誤差、評定者の違いによる誤差という3種類のランダム誤差の程度を総合的に把握することが可能となり、測定尺度としての信頼性を正確に評価することができる。

第三に、尺度の妥当性について、年齢・月齢による推移、因子構造、基準関連妥当性、臨床群のスコアプロフィールという3点から検討した。VABS-IIは適応行動の発達を評価する尺度であるため、その得点は年齢・月齢とともに徐々に上昇していくことが想定される。この想定を検証するため、年齢・月齢にともなう得点の推移を検討した。因子構造については、原版において想定される二次因子構造が日本人のデータにも適用可能か否かを確認的因子分析によって検証した。基準関連妥当性については、国内で開発・標準化された適応行動の測定尺度である旭出式社会適応スキル検査、国際的に広く利用される不適応問題の測定尺度である Child Behavior Check List (CBCL)、適応行動と関連が深い知的能力 (IQ) および自閉症特性 (広汎性発達障害日本自閉症協会評定尺度; PARS) との関連から多角的に検討した。さらに、臨床群として、知的障害、自閉症スペクトラム障害 (ASD)、ADHD、視聴覚障害のデータを収集し、各群のスコアプロフィールから、

VABS-II が様々な障害による適応上の困難を的確に把握しうるかを検討した。

B. 方法

1. VABS-II 日本版の開発

本事業が開始される前年度の 2008 年度より、VABS-II の翻訳と日本版の開発を開始した。まず分担研究者の黒田および研究協力者の田中によって原文を忠実に翻訳した翻訳原案を作成した。この原案をもとに、研究班全体で訳語の確認や生活習慣や文化の合わない記述について、原版と照らし合わせながら確認を行い改訂していった。また、日本において必要と考えられる事柄については新たに項目を追加した。原版からの主な変更点を Table 1 に示す。文法、文字などの言語形態の違いによる変更が大部分であった。また、下位領域身辺自立では、スプーン・フォーク・ナイフに関する項目を、箸に関する項目に変更した。その他、項目内の人名、貨幣などの記述も日本のものに変更した。

こうして作成された VABS-II の日本版について、バックトランスレーションを行い、原著者の確認及び使用許諾を得た。また、後述するように、VABS-II は各下位領域の中で比較的獲得の容易な行動から難しい行動へと、項目の難易度に沿った項目順となっている。この項目順についても、日本の調査における測定値に基づいて再検討を行う必要があると考えられたため、項目順の再配置について原著者に確認と承諾を得た。

本研究では、このようなプロセスを経て開発された VABS-II 日本版を用いて調査を実施した。

2. 対象者

本研究では、VABS-II 日本版の標準化サンプルとして一般群 1367 名のデータを収集した。また、尺度の信頼性・妥当性検証のために一般群延べ 130 名と発達障害・知的障害の診断を受けた臨床群 259 名のデータを収集した。

一般群については、地域の偏りが少なくなるよう全国 28 都道府県において調査を実施した。調査の開始前に以下の 3 項目について尋ね、全てが「いいえ」であるもののみを対象に面接を実施した。

- ① これまでに頭部外傷を受けて、意識がなくなったことがありますか？
- ② これまでに学習障害、ADHD（注意欠陥多動性障害）、情緒障害などを指摘されたことがありますか？
- ③ これまでに神経の病気をしたことがありますか？（例：てんかんなど）

臨床群については、全国 28 都道府県の医療・心理・教育機関を受診し、熟練した精神科医により DSM-IV の診断基準に基づいて何らかの発達障害（ASD、ADHD、LD）または知的障害の診断を受けた 259 名が調査の対象者となった。

一般群および臨床群のサンプルに関する詳細は結果の項に後述する。

3. 回答者

本研究で日本版の開発を行う VABS-II の Survey Interview Form（面接調査フォーム）の回答者には、対象者の日常をよく知っている成人が選ばれる。対象者が若年の場合は両親、養育者などの保護者が回答者

となるケースが多い。また、対象者が成人の場合には、配偶者、成人の家族などが考えられる。対象者が家族と同居していない場合には、対象者の日常を観察できる立場にいる人物（施設職員、カウンセラー、職場の上司など）が回答者に選ばれることもある。本研究においても、同様の基準を満たす回答者からの回答を得た。

4. 実施方法

VABS-II の面接調査フォームでは、半構造化面接のスタイルをとっている。これは、記録用紙に書いてある各項目をそのまま読んで質問していくのではなく、全体的な会話から次第に詳細事項へと移行するという流れで、自然な会話の中で各項目について質問していく方法である。不適応行動領域を除いて、各下位領域の項目は適応行動の発達段階に沿って並んでいるが、質問項目が多少前後しても構わないとされている。また、ひとつの行動がその程度によって複数の質問項目に分けられている場合（例えば、15 分話を聞く、30 分話を聞くなど）、一つの話題の中で複数の項目をスコアすることも可能になる。さらに、半構造化面接法では、さまざまな形で回答者に質問していくことによって、対象者のより正確な情報を聞き出すことが出来る。面接時間はマニュアルでは 20～60 分とされているが、対象者の年齢、障害のコンディションなどによって面接時間がそれより長くなることもある。

調査を実施する面接者に対しては、1 時間～2 時間程度の研修会を設け、面接の方法や、データの処理の仕方、個人情報の保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し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 (安達他, 2006; 神尾他, 2006; 辻井他, 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、VABS-II 日本版の適応行動尺度および不適応行動尺度について、項目反応理論 (IRT) に基づく項目分析とテスト情報量の評価を行った。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、内的整合性、再検査信頼性、評定者間信頼性の 3 つの観点から検討した。

最後に、VABS-II の妥当性について、因子構造、年齢にともなう得点推移、基準関連妥当性、臨床群のスコアプロファイルという 4 つの観点から検討を行った。

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) 項目分析

VABS-II 日本版の適応行動尺度および不適応行動尺度の各項目について、項目分析を行った。

適応行動尺度については項目反応理論(IRT)に基づく項目パラメータの推定を行った。IRTは、評価項目群への反応に基づいて、対象者の特性や評価項目の困難度・適合度を推定するためのテスト理論である。IRTは、項目への回答(VABS-IIでは0、1、2の3段階)という離散的な結果から、確率論的に項目の困難度を求めるため、素点に基づく分析よりも厳密な困難度の評価を可能にする。前述のとおり、VABS-II日本版の簡便な実施を可能にする打ち切りルールを適応行動尺度に適用するためには、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ため、原版と同様、本研究においてもIRTによって厳密な困難度の評価を行う。

IRTには、困難度のみを項目パラメータとして推定する1パラメータロジスティックモデル(ラッシュモデル)、困難度と識別力を推定する2パラメータロジスティックモデル、さらに偶然による正解の確率を考慮する3パラメータロジスティックモデルの3種類が存在するが、本研究では、主に困難度を問題にしていること、原版においてもラッシュモデルが採用されていること

から、ラッシュモデルを用いて検討を行う。また、2段階で回答を求める場合は2値型IRTモデルを用いるが、ここでは3段階で回答が求められるため、多値型IRTモデルである部分採点モデル(partial credit model)を用いた。原版において、分析プログラムはWINSTEPS(Linacre, 2003)が使用されているため、ここではWINSTEPS(Linacre, 2012)を使用し、同時最尤推定法によるパラメータ推定を行う。

IRTに基づく分析では、項目の困難度に加えて、項目適合度が算出される。これは、様々な個人の能力値における項目得点の期待値と実際の観測値のズレを評価する指標であり、古典的テスト理論における項目合計相関(各項目の得点と尺度得点の相関。項目が尺度の内的整合性に寄与するか否かを評価するための指標)と同様の機能を果たす。ラッシュモデルの適合度指標には、外れ値の影響を受けやすいoutfitと、外れ値よりも全体的な反応のパターンにセンシティブなinfitの2つの指標が存在する。いずれの指標も0.5から1.5までの範囲が望ましく、1.5よりも値が大きくなると、モデルとデータの適合度が低く、項目が尺度の信頼性に寄与しない、または悪影響を及ぼすことを示す(項目合計相関が低い状態に対応する)。逆に、0.5より値が小さくなると、モデルとデータの適合度が高すぎ、項目がその尺度において冗長であることを示す。

Table 4~Table 14に、適応行動尺度の各下位領域の項目得点の記述統計とラッシュモデルに基づいて算出された困難度・項目適合度を示す。先述のとおり、打ち切りルールを適用するためには適応行動尺度のいずれの領域においても、項目の順序が後にな

るほど平均得点が下がり、困難度が上昇することが望ましい。全ての尺度において、これらの傾向が認められたが、部分的には平均得点や困難度の並びが項目順と一致しない箇所も見られた。そこで、最終的な項目番号は困難度パラメータが項目順に上昇するように入れ替えを行い、配置した。表中の番号は入れ替えを行った後のものであり、調査時の番号は旧と表記した。例えば、下位領域「受容言語」において、調査時の項目 6 はその後の項目 7、項目 8、項目 9、項目 10 よりも困難度が高いため、それらよりも後ろに配置された。逆に、項目 10 は困難度が低いため前に配置されている。

Table 15 に不適応行動尺度の各項目の平均・標準偏差と修正済み項目一合計相関を示す。いずれの項目も 0.01 から 0.21 という低い平均値を示している。このことは、不適応尺度の各項目が一般健常者に広く見られる行動特徴ではなく、発達障害や精神疾患に関連する特異的な不適応行動を捉えていることを反映している。各項目の項目一合計相関は、下位尺度単位、尺度全体のいずれにおいても十分な値が示されている。

3) テスト情報量

IRT 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。

4) 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。

5) 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。

6) 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、全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) 内的整合性

適応行動尺度の各下位領域・領域・尺度全体における年齢帯ごとの α 係数をTable 16に示す。 α 係数は、内的整合性の観点から尺度の信頼性を評価するための指標であり、一般に.70以上の値を示すことが望ましいとされる。この値は打ち切りルールを適用することによって高くなることが想定されるため、ここでは打ち切りルールを適用する前の評定値を用いて算出した。

ほとんどが.70以上の値を示しているが、19-21歳の受容言語尺度など、高年齢帯を中心に低い値を示すものがある。これらはテスト情報量の低い範囲に分布する年齢帯と一致しており、尺度の内的整合性の低さによるものではなく、対象者の能力が高いため個人差が測定できていない可能性を示している。VABS-IIは障害児者の適応状況の把握を目的とする尺度であるため、一般群の成人を対象とした場合はこのような結果が得られると考えられる。なお、これは原版においてもこれは見られる傾向である。

次に不適応行動尺度の各下位尺度および総得点の年齢帯ごとの α 係数をTable 17に示す。「不適応その他」については、下位尺度単位で評価を行わないこととなっているため、 α 係数を算出していない。

下位尺度単位では、特に「不適応内在化」において一部低い値を示すものもあるが、概ね.70を超える値が得られている。尺度全体（不適応内在化+不適応外在化+不適応その他）では、全ての年齢帯で.70を越える値が示された。

2) 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。

3) 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) 因子的妥当性

VABS-II では、11 の下位領域が 4 つの領域によって規定され、さらにそれらが適応行動という単一の概念によって規定されるという二次因子構造が想定されている。ここでは、標準化サンプルを用いてこの二次因子モデルの妥当性を検討した。

まず、4 つの年齢帯ごとに下位領域の標準得点間の相関を算出した (Table 20～Table 23)。全般的にそれほど顕著な差はないが、同一領域内の下位領域間では他領域の下位領域との相関よりもやや強い相関が見られる。ただし、年齢帯によって相関の表れ方には若干の差が見られる。

より明確に因子モデルの検証を行うため、確認的因子分析による検討を行った。二次因子モデルの対立仮説として、単一の因子によって 11 の下位領域が規定されるという一次因子モデルと全ての下位領域が独立である (相関を持たない) という独立モデルを想定し、適合度を比較検証した。観測変数には各下位領域の標準得点を用いた。

Table 24 に各モデルの適合度を示す。 χ^2 値の差に基づく検定の結果、いずれの年齢帯においても、独立モデルより一次因子モデルが、一次因子モデルより二次因子モデルが有意に良好な適合を示した。また、CFI、RMSEA、AIC の 3 指標でも、全ての年齢帯

で二次因子モデルが最も良好な適合を示した (ただし、0-2 歳では RESEA で一次因子モデルが二次因子モデルより良好な値を示した)。

二次因子モデルにおける CFI (.90 を超えると適合が良好とされる) の値は、16-92 歳で .851 というやや低い値が示された他は、いずれの年齢帯も .90 を上回っており、良好な適合が示された。RMSEA (.05 以下が良好、.05～.10 がやや良好) は、16-92 歳で .119 というやや高い値が見られたが、他の年齢帯はいずれも .08 前後の値を示しており、許容しうる範囲の適合が示された。

Figure 35～Figure 38 に二次因子モデルにおける各下位領域の負荷量を示した。各下位領域の一次因子 (領域) に対する負荷量は、0-2 歳、3-6 歳の粗大運動や 7-15 歳の受容言語がやや低い値を示したが、全体的には中程度以上の負荷量を示すものが大部分であり、各下位領域が一次因子の指標として機能することが示された。また、一次因子の二次因子 (適応行動) に対する負荷量はいずれも .80 以上の高い値を示し、一次因子間の関連が二次因子によって説明されることが示された。

以上の結果から、原版で想定される二次因子構造が VABS-II 日本版にも許容しうる範囲の適合性を有することが示された。

2) 年齢および月齢にともなう得点推移

VABS-II は適応行動の発達を評価するツールであるため、その素得点は対象者の年齢にともなって上昇すると予測される。そこで、Figure 39～Figure 49 に適応行動尺度の各下位領域の素得点の年齢にともなう平均値の推移を標準偏差とともに示す。い

れの下位領域においても年齢にともない滑らかに得点が上昇しており、適応行動尺度の妥当性が示されている。また、地域生活、遊びと余暇、粗大運動などの下位領域では、50歳以降で得点が低下する傾向が見られている。このことは、当尺度が行動の発達のみならず老化に伴う衰えをも捉えるツールであることを示している。

各下位領域の得点の勾配に着目すると、下位領域によって発達の初期に顕著な勾配が見られるものと、成人期まで長く勾配が続く下位領域がある。例えば、コミュニケーション領域の受容言語や表出言語は0歳から5歳ごろにかけて大きな勾配が見られるが、その後は大きな変化を示さない。一方、日常生活スキル領域の家事や地域生活は、5歳から7歳に最も大きな上昇を見せ、さらに成人後まで勾配が続き、標準偏差も一定の値を保ち続ける。VABS-IIの各下位領域は、互いに異なる年代の適応行動の発達にセンシティブであり、それらを組み合わせることで全年代の適応行動の包括的な評価が可能になっていることが、こうした結果からも示されている。

また、VABS-IIは、低年齢ほどきめ細かく発達を評価できるよう構成されている。そこで、0～23ヶ月までの月齢ごとの下位領域得点の平均値の推移を標準偏差とともにFigure 50～Figure 59に示した（読み書きは2歳以下は評価しないためグラフを示していない）。いずれの下位領域においても、月齢にともないおおむね滑らかに得点が上昇する様子が見取れ、当尺度が月単位での発達評価にも利用可能であることが示された。発達初期は、自閉症などの発達障害の早期発見において重要性の高い時期であ

り、その時期において月単位での発達評価を可能にするVABS-IIはきわめて有用性が高いと言える。得点の勾配と標準偏差に着目すると、全体に11ヵ月前後から得点の上昇幅が大きくなり、標準偏差も大きくなる傾向が見取れる。この時期は、運動、言語、因果性理解、意図理解など、様々な側面での発達が急速に進む時期であり、それと同時に、発達の個人差が顕在化する時期でもある。VABS-IIは、こうした発達の実情を明確に捉えていることがわかる。

全ての下位尺度において、年齢・月齢のいずれについても成長に伴う上昇が見られる一方、領域によっては成人期後期より老化に伴う下降が見られることはVABS-IIの妥当性を示している。

3) 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関連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。

4) 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て実施することで、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。

5) 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 の関連を検討した。

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、不適応行動との関連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

項目分析の結果から、VABS-II を構成する 400 以上の項目の全てが、属する下位領域に対して十分な適合を有することが明らかになった。VABS-II の各項目はもともとアメリカで開発されたが、日本の言語体系や文化に合うように改訂を行って日本版を開発したことで、日本人の対象者に対しても有効に機能するものになったことが示唆される。

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

内的整合性の観点から各下位領域の信頼性を検討したところ、高年齢帯で一部に低い α 係数が見られた以外は、ほぼ十分な値が得られた。一部の年齢帯で低い α 係数が示された下位領域は、いずれもその年齢帯が示す特性値におけるテスト情報量の値が低い値を示しており (Figure 1~Figure 11 参照)、尺度そのものの内的整合性が低いというよりも、当該年齢帯の対象者が尺度によって個人差を検出しようとする範囲を超える能力を有していたために α 係数が低い値を示したと推察される。しかし、これは尺度の上限付近の高得点を示す一般群の高年齢帯の対象者に特有の結果であり、実際に VABS-II のアセスメントの対象になる障害児者においては十分な信頼性をもって評価を行うことができると考えられる。

再検査信頼性および評定者間信頼性についても検討を行った結果、一部の下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、下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 日本版の妥当性について、まず因子構造の観点から検討した。原版では11の下位領域が4つの一次因子（領域）によって規定され、さらにそれらが1つの二次因子（適応行動）によって説明されることが想定されている。この二次因子モデルがVABS-II 日本版にも適合するか否かを検討するため、確認的因子分析による検証を行った。対立仮説として、11の下位領域が直接1つの適応行動概念によって規定されるという一次因子モデルと、全ての下位領域を独立と見なす独立モデルを想定し、比較検証を行った。その結果、いずれの年齢帯においても二次因子モデルが他の2モデルよりも良好な適合を示し、二次因子モデルの妥当性が示された。

VABS-II は、知能検査が知的能力における発達の程度を測定するのと同様に、適応行動における発達の程度を測定する尺度である。言い換えれば、年齢段階に応じた適応行動を獲得できているかを評価することがVABS-IIの基本的な目的である。そのため、その測定値は年齢とともに徐々に上昇することが想定される。そこで年齢とともに素得点の推移を検討した結果、いずれの下位領域においても、年齢とともに平均値が滑らかに上昇することが示された。また、その勾配の表れ方は下位領域によって異なり、例えば受容言語では5歳ごろにピークを迎え、その後一定の値を維持するが、表出言語では8歳ごろ、読み書きでは成人後まで得点の上昇が続いていた。このことから、VABS-II 日本版が適応行動の発達を鋭敏に捉える尺度であること、また、互いに異なる年代の発達に敏感な複数の下

位領域の組み合わせによって全年代の発達に対応していることが明らかになった。

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、下位尺度単位でも概念的に関連の深い下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・下位領域の収束的・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.

井瀬知美・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
総合研究報告書に記載
- H. 知的財産権の出願・登録状況
特になし