

平均値 - 5×標準偏差 > から < 平均値 + 3×標準偏差 > の範囲に対応するとされている。したがって、v-scale は平均値が 15、標準偏差が 3 の（原則的には）正規分布する標準得点（standard score）として表示される。

もし、年齢、月齢ごとに十分な数（できれば 1,000 名以上）のデータがあれば、このような変換は容易である。すなわち、各段階のデータの（たとえば）累積度数分布から、上記のような条件を満たすように得点を変換することは技術的にさほど困難ではない。ただし、すでに随所で示されたように、多くの領域で素点はある程度以上の年齢段階では頭打ちになることが多く、一定の問題が残るのは事実である。

しかしながら、現実には本研究において、対象者数はすべての年齢段階を通して、1400 に達しない。オリジナルの VABS-II においても、標準化に用いられた対象者数は 3000 強にすぎない。したがって、広範囲の、かつ、相当に細分化された年齢、あるいは月齢段階ごとに v-scale を求めるためには、一定の工夫が必要となる。

オリジナル版において用いられた手続きは以下の通りである。

- (1) 全対象者を 23 のグループに分ける（各段階に含まれる対象者数は最大でも 200、最小で 50 を少し超える程度である）。
- (2) グループごとに、素点の平均値、標準偏差、歪度、尖度を計算する。
- (3) 平均値と標準偏差を年齢に対してプロットし、滑らかな曲線になるように平滑化（smoothing）を行う。
- (4) Johnson (1947) による曲線の当てはめのプログラム、23 の年齢段階ごとに、上

で求めた、平均値、標準偏差、歪度、尖度を入力して、素点を正規分布に変換する。

- (5) 得られた得点を 3 倍して 15 を加えることにより v-scale に換算する。

この方法には、いくつかの問題点が指摘される。まず、比較的少数の標本から計算される平均値や標準偏差には、当然標本誤差が含まれるから、それを平滑化によって補正することは許容される手続きであろうが、計算過程にデータの 3 乗と 4 乗が含まれる歪度や尖度の方が、標本誤差は大きいはずであるにもかかわらず、Johnson の方法への入力としてそれらがそのまま用いられるのは疑問であり、その段階でかなりの誤差、あるいは偏りを発生する可能性がある。偏りのチェックのために、（年齢段階ごとに）パーセンタイル順位から正規分布への変換を行なったものと比較が行われたりしているが、最終的な補正にあたっては、かなり主観的な処理が行われているようである。

こうした処理は、オリジナルの約 3 分の 1 のデータ数しかない今回のデータに適用するとなると、より大きな偏りや誤差を有無心配がある。

以上のような純技術的な問題以外にも、こうしたやり方では、せっかく 1 次元の連続体として得られた素点を、年齢段階を超えて比較することが簡単にはできなくなる。実際、ある対象者が、同年齢の標準と比較して平均値より 2×標準偏差だけ遅れていることが分かったとしても、それがどの程度の年齢段階に対応するかを即座には判断できなくなってしまふのである。したがって、今回のデータにより適した、v-scale の計算方法を考える必要が生じたわけである。

B. 問題点の整理

改めて、我々が直面している問題について整理しておこう。

まず、我々は、著しく年齢によって異なった分布をもつ複数の尺度得点をもっている。図1は、「身辺自立」の月齢0~6箇月(0:0-0:6と表記)年齢2~4歳(2:0-4:0), 10~14歳(10:00-14:00)の素点の分布を示したものである。ここで評価されている項目が示しているタスクについて、6箇月未満の乳児はほとんど遂行できず、10歳をこえた児童・生徒はほとんど遂行することができ、ともに極めて分散が小さい、他方、2歳から4歳の間の幼児については、遂行できるかできないかの幅が極めて大きく、得点の分散も極めて大きい。

われわれの解決しなければならない課題は、この図の横軸である素点の尺度を適切に変換し、図2のような正規分布に変換することである。このように変換された尺度を、改めて年齢段階ごとに平均値が15、標準偏差が3の標準得点に変換すれば、(原理的には)目的とする *v-scale* を得ることができるはずである。

C. 全年齢段階が正規分布する変数への変換

こうした変換を行うために、我々が手にしているデータの数はあまりに少ない。したがって、理論的には多少とも不十分であっても、頑健な方法を用いる必要がある。

まず、図1のような度数分布を直接扱うことによってこの変換を行うことは困難であるように思われるので、図3に示すような累積度数分布を用いることを考える。こ

こでは、表1のように、年齢・月齢別に13のグループに分割している。

累積度数分布は、度数が0の区間においても、関数の値は一定となるだけで0となることはない。また、単調非減少関数であり、急激な増減のない安定した値を保つことから、少数のデータから結果を求めなければならないわれわれの状況には適している。ただし、各点の数値は、直前までの値に条件づけられた、独立でないものであるから、確率論的に正確な扱いには必ずしも適していないという欠点も指摘しておかなければならない。

これが、図4のように正規分布の累積曲線(分布関数と腰部のが普通)のようになるような横軸の変換方法が見つかれば、目的はとりあえず達成されたことになる。

D. 系列範疇法の利用

このようなことを可能にする方法として、系列範疇法(*method of successive categories*)と呼ばれる方法がある。この方法を、ここで分析の対象としているデータの形にあわせて説明しよう。ある領域の素点の満点が k であるとし、変換後の尺度上で隣接する2つの素点の境界が t_g ($g = 0, \dots, k-1$) という値をとるものとする。また、グループ i ($i = 1, \dots, n$) が平均値 m_i 、標準偏差 s_i で変換された尺度上で分布しているとして、グループ i において得点 g 以下の点をとる対象者の割合(相対累積度数)を標準正規分布に変換した値を z_{ig} とする。このとき、図4のように、すべてのグループが正規分布するとすれば、

$$z_{ig} = \frac{m_i - t_g}{s_i} \quad (1)$$

が成立するはずである。Diederich, Messick, & Tucker (1957) は、この方法の最小2乗解を提案しており、ここではそれを利用することとした。

なお、この方法については、さらに進んだ最尤法等による解も得られているが、十分な数のデータがない場合、一般には最尤法より最小2乗法の方がアルゴリズムの安定性が高く、あえて最小2乗解を利用することとした。

E. さらに細分化した月齢・年齢段階の尺度値

図5に、この方法を適用して得られた、素点と尺度値 (t_g) の関係を示した。また各グループの分布は、既に図4に示したようになった。この縦軸に示した尺度値上で、すべての年齢段階の個体の得点を比較することができ、また、各グループの v-scale を (平均値が 15, 標準偏差 3 の) 標準得点として算出できる。

実際の使用に供するには、表1のグループ分けでは粗すぎるので、より細かいグループングが必要になる。これについては、当面、表1の各群の平均値と標準偏差 (図6) を直線補間することによって中間段階の平均値と標準偏差を求めるといった簡便な方法をとっている。

F. 残された問題

ここで述べたやり方は、とりあえず大部分のケースにおいて、一応受け入れ可能な結果を生み出したように思われる。ただし、解決すべきいくつかの問題も残されている。中でも大きな問題は、ある1つの領域 (「表

出言語」) において、オリジナルとかなりかけ離れた結果が生じたことである。具体的には、低い年齢段階の間の平均差とその近辺の段階の標準偏差が小さすぎるのである。それば、ここで述べた方法が、図3や図4に見られるように、直接比較可能できないほど離れた区間の間での尺度値を、(1) の数理モデルを通して比較可能なものとしようとしているメカニズムに、時に無理が生じるせいではないかと考えられる。結局現状では、そうした場合に A で退けた平滑化を援用しなければならなくなっている (Luke, 2008)。今後は、より詳細にデータの特徴を検討しながら、より合理的で受け入れ可能な買いを得るための方法を探求していく必要がある。そのための里程標といった意味で、現状報告をしておきたい。

G. 引用文献

Diederich, G.W., Messick, S.J. & Tucker, L.R. (1957). A general least squares solution for successive intervals. *Psychometrika*, **22**, 159-173.

Luke, K. (2008). *Semiparametric regression for social sciences*. Chichester: Wiley.

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

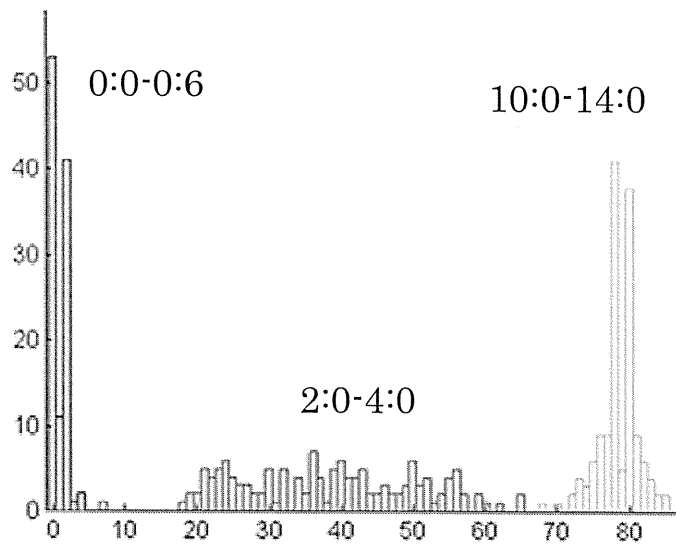


図1 「身辺自立」の年齢別素点の分布

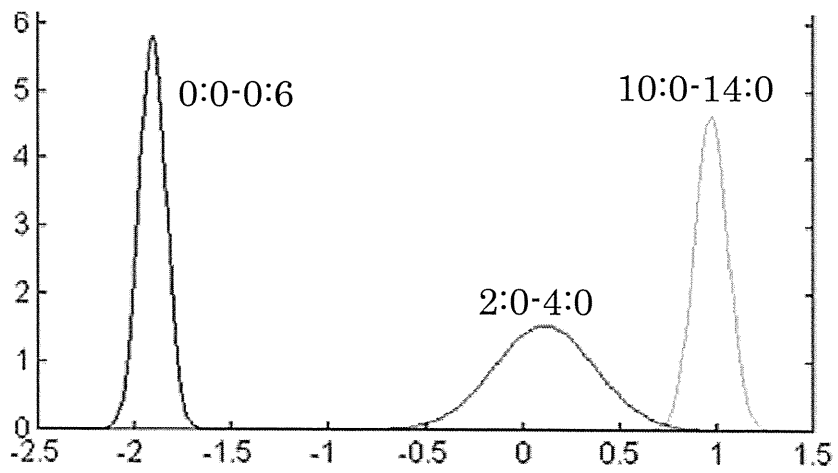


図2 各年齢段階が正規分布する共通尺度

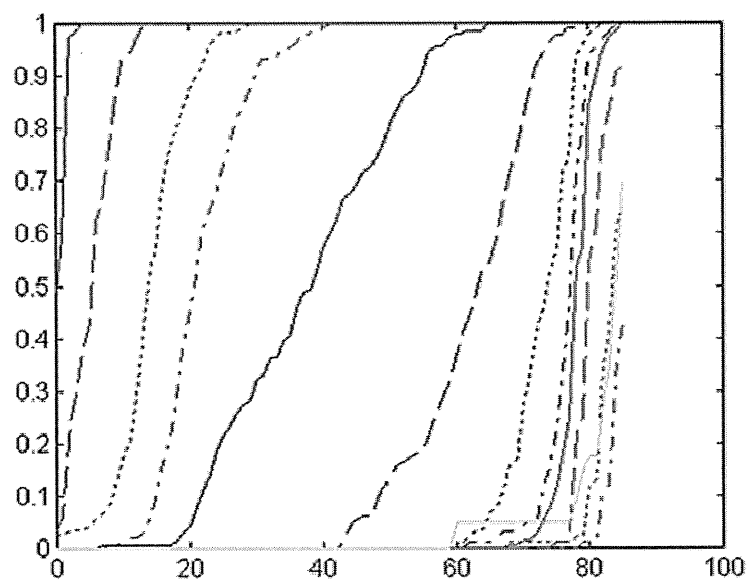


図3 「身辺自立」の年齢段階別の累積度数分布

表1 対象者のグループと人数

グループ	人数
0:0-0:6	108
0:6-1:0	114
1:0-1:6	117
1:6-2:0	107
2:0-4:0	143
4:0-6:0	132
6:0-8:0	93
8:0-10:0	92
10:0-14:0	146
14:0-19:0	100
19:0-32:0	83
32:0-72:0	92
72:0-92:0	40

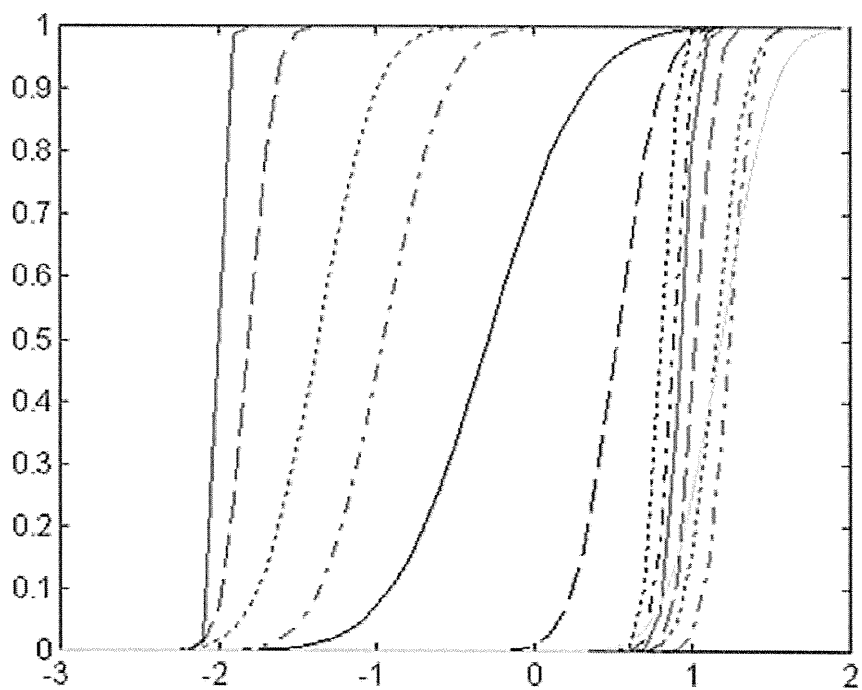


図4 「身辺自立」の年齢別分布関数（推定値）

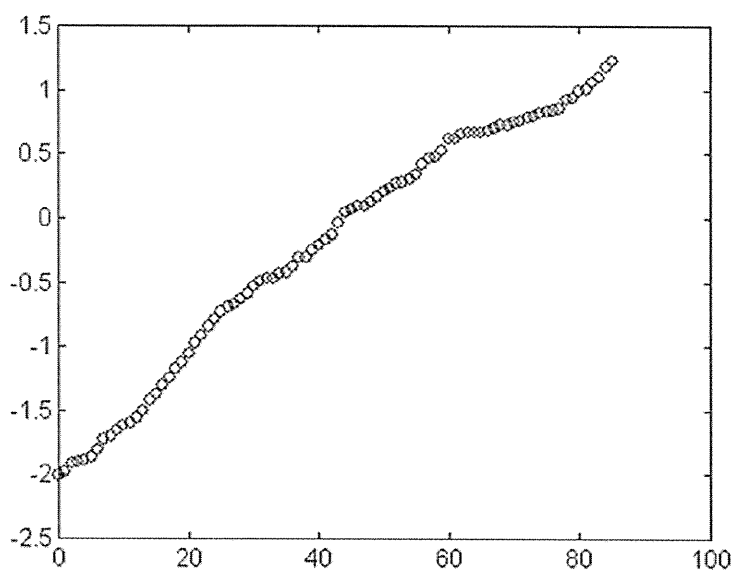


図5 「身辺自立」の素点（横軸）と尺度値

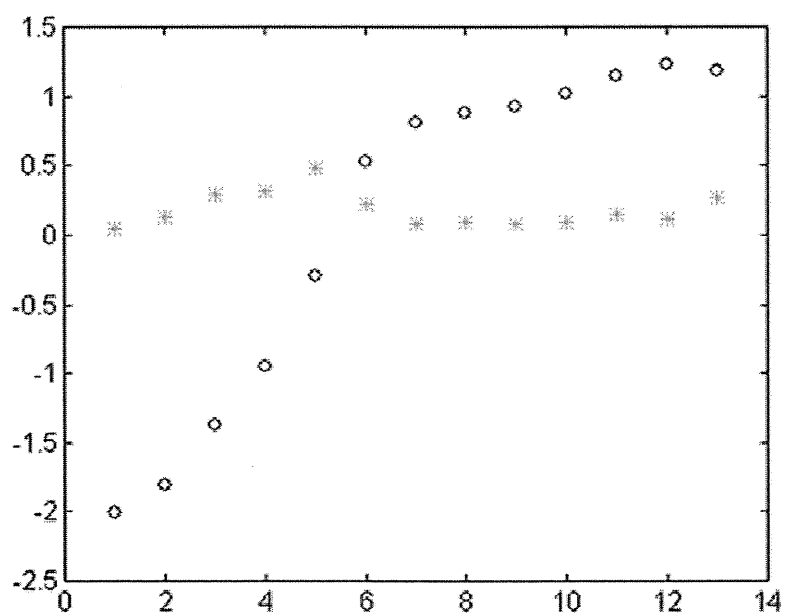


図6 表1のグループの「身辺自立」の平均値(○)と標準偏差(*)

感覚プロフィール日本版の標準化と信頼性・妥当性の検証

分担研究者 萩原 拓 北海道教育大学旭川校
分担研究者 岩永竜一郎 長崎大学大学院医歯薬学総合研究科
研究協力者 平島太郎 名古屋大学大学院教育発達科学研究科
分担研究者 伊藤大幸 浜松医科大学子どものこころの発達研究センター
研究代表者 辻井正次 中京大学現代社会学部

研究要旨

発達障害児者の多くは感覚刺激への反応異常を示すことが報告されている。感覚刺激への反応異常は、日常生活の行動にも影響することが示唆されており、その有無や程度を評定することは支援の際に不可欠である。ところが本邦においては、それを的確にアセスメントするツールが不足していた。従来から使われてきた感覚チェックリストは存在するが、対象年齢が幼児に縛られており用途が限られていた。そして、海外で使われているツールとは異なるために検査結果を先行研究に照らし合わせ、支援を考えることや研究に用いることができなかった。

そこで本研究は、感覚刺激に対する反応異常のアセスメントツールとして国際的に広く用いられている感覚プロフィールの日本版を開発することとした。初年度は、感覚プロフィールの乳幼児版、子ども版、青年成人版を翻訳し、その翻訳精度を確認した。第二年度は自閉症スペクトラム障害（ASD）児者や知的障害（MR）児者のデータを収集し、その傾向を分析した。最終年度にあたる本年度は、一般群延べ 2888 名のデータを収集し、感覚プロフィール日本版の標準化を行うとともに、ASD 児者や MR 児者のデータとの比較から、尺度の妥当性検証を行った。その結果、感覚プロフィール日本版の 3 バージョンに関して、乳幼児から成人に至るまでの各年齢段階における標準値が得られた。また、各バージョンは、いずれも十分な内的整合性を有すること、一般群と ASD 群や MR 群との間で顕著な差異が見られること、知的水準や自閉的特性の程度と相関があること、日常生活への適応に必要となる適応行動や適応を阻害する不適応行動と関連があることが示され、尺度としての信頼性・妥当性が確認された。

A. 研究目的

発達障害児者には感覚刺激に対する反応異常が見られることが多い。Gomes et al. (2008) の自閉症スペクトラム (ASD) 児の感覚の問題についてまとめた研究レビューではその 90%に感覚過敏が見られることがわかっている。また、Tomchek & Dunn (2007) の研究で自閉症児の 95%は、感覚処理の問題をとらえる尺度、短縮版感覚プロフィール (Short Sensory Profile) (Dunn, 1999) の総合点が平均値より-1SD以下であり、その多くが感覚処理の問題を持っていることが示されている。

感覚の問題は対人関係、情動、行動などに見られる問題と関係していることが知られている。Hilton et al. (2010) は高機能 ASD 児の感覚刺激に対する反応異常が社会性の障害の重症度と有意に相関が見られたことを報告している。Lane et al. (2010) は注意欠陥多動性障害 (ADHD) 児、定型発達児共に感覚過敏と不安の関係が認められたことを報告している。ASD のある当事者の自伝からも感覚の問題によって情動が乱され、日常生活がままならない状態になってしまう実情が記述されている (グランディン&スカリアーノ, 1994; ウィリアムズ, 2000)。このように発達障害児者の感覚の問題は学校や社会生活での適応に影響することが多い。

そのため、発達障害児の感覚の問題を把握し対応を検討する必要がある。しかしながら、本邦ではそれを的確に評定するツールが不足していた。これまで、我が国では感覚面の評価に日本版感覚インベントリー (Japanese Sensory Inventory Revised: JSI-R; 太田, 2002; 太田, 2004) が用いられることが多かったが、これは4-6歳児を対象とし

た評価であるため、それ以外の年齢に適用できなかった。また、海外で使われているツールとは異なるために検査結果を先行研究に照らし合わせ、支援を考えることや研究に用いることができなかった。一方、海外では感覚刺激に対する反応の評定には Sensory Profile (SP) (Dunn, 1999; Dunn, 2002; Brown & Dunn, 2002) が用いられることが多い。SP は発達障害児者などの感覚刺激に対する反応異常を評定する質問紙式の検査である。この検査はアメリカで標準化されており、信頼性、妥当性共に問題がないことがわかっている。SP には対象児者の年齢に応じて、乳幼児用 SP (0-36 ヶ月; 以下 SPI)、SP (3-10 歳用)、青年期・成人用 SP (11 歳以上用; 以下 SPA) の3バージョンが標準化されている。SP を活用すれば、発達障害児者の感覚刺激に対する反応異常についての信頼性のあるデータが収集でき、その結果と適応行動の問題をとらえることができると考えられる。そして、臨床・教育現場での指導、家族指導において重要な情報を提供することになる。以上のように SP を本邦において再標準化し、実用化することは急務である。

そこで、本研究班では、これら3バージョンの SP の標準化を目指して、大規模な調査を行ってきた。初年度は、各バージョンを翻訳し、バックトランスレーションによって翻訳精度を確認した。第二年度は ASD 児者や知的障害 (MR) 児者のデータを収集し、その傾向を分析した。最終年度にあたる本年度は、一般群 2888 名のデータを収集し、感覚プロフィール日本版の標準化を行うとともに、ASD 児者や MR 児者のデータとの比較から、尺度の妥当性検証を行った。

B. 方法

1. 対象

本研究では、3バージョンの日本版 SP の標準化のために一般群延べ 2888 名のデータを収集するとともに、尺度の妥当性検証のために発達障害・知的障害の診断を受けた臨床群延べ 490 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、意識がなくなりましたか？
- 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者が調査の対象者となった。

調査にあたっては、対象者の保護者、配偶者、きょうだい、施設職員等、対象者をよく知る者に面接および質問紙による調査を実施した。ただし、11 歳以上の対象者については本人からも SPA について回答を得た。

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。

2. 調査内容

1) SPI（乳幼児版 SP）

36 ヶ月までの乳幼児を対象とする SPI は、0-6 ヶ月児を対象とする 36 項目と 7-36 ヶ月児を対象とする 48 項目からなる。0-6 ヶ月用は、一般、聴覚、視覚、触覚、前庭覚の 5 セクション、7-36 ヶ月用はこれに口腔感覚を加えた 6 セクションからなる。また、セクションとは別の概念的区分として、感覚探究、感覚過敏、低登録、感覚回避の 4 象限が存在する。これら 4 象限は、行動反応／自己制御の次元と神経学的閾値の次元によって区分されたもので、積極的反応・高閾値が感覚探究、消極的反応・低閾値が感覚過敏、消極的反応・高閾値が低登録、積極的反応・低閾値が感覚回避にそれぞれ対応する。各項目は日常の感覚経験に対する子どもの反応を記述している。保護者はこれらの項目に描かれた反応を子どもが示す頻度を 5 段階（1. いつも、2. しばしば、3. ときどき、4. まれに、5. しない）で回答する。評定値が高いほど、頻度が低いことを意味する。

スコアリングはセクションごとおよび象限ごとに合計点を算出した後、それらを標準分布に照らし合わせて行われる（ただし、0-6 ヶ月用はセクションごとのスコアリングは行わない）。0-6 ヶ月用については、各象限の得点が平均±1SD 以内の場合は「一般的反応」、平均+1SD 以上または平均-1SD 以下の場合は「相談とフォローアップ」の対象と判断される。7-36 ヶ月用については、各セクション・象限の得点が平均±1SD 以内の場合は「一般的反応」、平均+1SD～2SD または平均-1SD～2SD の場合は「異なる可

能性」、平均+2SD以上または平均-2SD以下の場合は「明らかに異なる」と判断される。

本研究では、0-36ヶ月の対象者の保護者にSPIへの回答を求めた。

2) SP (児童版 SP)

3-10歳の子どもを対象とするSPは125項目によって構成される。質問項目は、「感覚処理」に関する6セクション（聴覚、視覚、前庭覚、触覚、複合感覚、口腔感覚）、「調整」に関する5セクション（耐久性・筋緊張、身体位置・動き、活動レベル、感情反応、感情・活動レベル）、「行為と感情」に関する3セクション（反応、感情的・社会的行為、閾値）の合計14セクションに分類されている。また、セクションとは別に、探索的主成分分析によって見出された9因子構造に基づく尺度区分も存在する（9因子区分では主成分分析で剰余項目となった44項目はスコアに含まれない）。対象児者の感覚刺激に対する反応を問う項目、例えば、「突然のまたは大きな音に拒否反応を示す」、「歯磨きなどの時に口に触られるのが嫌い」等によって構成されている。これに保護者が5段階回答（1. いつも、2. しばしば、3. ときどき、4. まれに、5. しない）をすることによって評定される。評定値が高いほど、頻度が低いことを意味する。

スコアリングは、14セクションごと、9因子ごとに合計点を算出した後、それを標準分布に照らし合わせて行われる。そして、平均-1SD以上を「一般的反応」、平均-1SD～-2SDを「異なる可能性」、平均-2SD以下を「明らかに異なる」と分類する。SPIおよびSPAとは異なり、得点が低い（反応性が高い）場合のみが異常判定の対象となる。

本研究では、3歳以上の対象者について、保護者、配偶者、きょうだい、施設職員等、対象者をよく知る者にSPへの回答を求めた。SPは原版では3-10歳の子どもが対象とされているが、青年成人版のSPAが自己評定形式を取っており、ASDやMRなどの発達障害を有する者には回答が難しいと考えられることから、本研究では11歳以上の対象者についてもSPを実施し、適用可能性を検討した。

3) SPA (青年成人版 SP)

11歳以上を対象とする自己評定形式のSPAは60項目から構成される。質問項目は味覚・嗅覚、運動、視覚、触覚、活動レベル、聴覚の6セクションからなるが、スコアリングに際しては、感覚探究、感覚過敏、低登録、感覚回避の4象限ごとに集計・評価される。SPIと同様、これら4象限は行動反応／自己制御の次元と神経学的閾値の次元によって区分され、積極的反応・高閾値が感覚探究、消極的反応・低閾値が感覚過敏、消極的反応・高閾値が低登録、積極的反応・低閾値が感覚回避にそれぞれ対応する。各項目は日常の経験に対する反応を記述したもので、そのような反応を示す頻度を本人が5段階（1. ほとんどしない、2. まれに、3. ときどき、4. しばしば、5. ほとんどいつも）で回答する。SPIやSPとは異なり、評定値が高いほど、頻度が高いことを意味する。

スコアリングは象限ごとに合計点を算出し、それらを標準分布に照らし合わせて行う。各象限の得点について平均-2SD以下を「一般よりかなり低い」、平均-2SD～-1SDを「一般より低い」、平均±1SD以内を「一

般と同程度」、平均+1SD~2SD を「一般より高い」、平均+2SD 以上を「一般よりかなり高い」と判定する。

本研究では、11 歳以上の対象者について、本人に SPA への回答を求めた。

4) 知能指数 (IQ)

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、185 名 (55.6%) について IQ の情報が得られた。

5) 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)、感受性 (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症状を記述したもので、「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。

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、対象者をよく知る者に対して PARS を実施した。

6) Vineland-II 適応行動尺度

Vineland-II 適応行動尺度 (Sparrow et al., 1984; Sparrow, et al., 2005; 以下、VABS-II と記載) は、様々な障害や疾患を抱える者の適応行動の発達や機能低下を評価するための半構造化面接形式の尺度であり、知的障害・発達障害のアセスメントを始め、国際的に幅広い研究・臨床の文脈で使用されている。本研究ではバックトランスレーションと約 1400 名のサンプルによる標準化のプロセスを経て開発された VABS-II 日本版を使用した。

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 領域 (コミュニケーション、日常生活スキル、社会性、運動スキル) 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、不適応内向、不適応外向、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。

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、全く行わないことを意味する。また、項目に記述された行動を示す機会がない場合にはN/O、項目に記述された行動をしているかどうか判断できない場合にはDKの評定が与えられる。N/OとDK(各下位領域3個以内)はスコアリングに際して1点と見なして計算される。ただし、各下位領域でDKが4つ以上つけられた場合には、その下位領域の得点は算出されない。

スコアリングにあたって、適応行動尺度については、94の年齢区分ごとの標準得点を算出し、分析に使用した。標準得点(15を平均、3を1SDとするスコア)は下位領域ごとに算出され、それらを4つの各領域内で平均することで領域得点を算出し、さらにそれらを平均することで総得点を算出した。不適応行動尺度については、不適応内在化および不適応外在化の素点の単純合計を下位尺度得点として使用した(原版に従い、不適応その他は下位尺度単位では使用しなかった)。総得点については、不適応内向、不適応外向、不適応その他の合計点を分析に使用した。

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、対象者をよく知る者に対してVABS-IIを実施した。

3. 分析

SPI、SP、SPAの3バージョンそれぞれに

ついて、以下の2点から検討した。

第一に、一般群サンプルに基づいて尺度の標準化のための分析を行った。まず尺度を構成する項目が有効に機能しているか否かを項目分析によって検討した。次に、尺度としての信頼性を内的整合性の側面から評価した。その上で、尺度得点の分布を確認し、原版と同様の年齢区分ごとの標準値(平均値、標準偏差)を算出した。

第二に、臨床群サンプルに基づいて尺度の妥当性を検証した。まず、診断(MR群、高機能ASD群、低機能ASD群)ごとの尺度得点の分布を確認するとともに、各群の平均値と一般群の標準値との差異を検討した。また、知的能力(IQ)、自閉的特性(PARS)、適応行動・不適応行動(VABS-II)との関連から、基準関連妥当性を検討した。ただし、SPIについては臨床群のデータが十分に収集できなかったため、検討を行わなかった。

C. 結果

1. SPI(乳幼児版)

サンプル

Table 1に参加者の内訳を示す。0ヶ月から36ヶ月の乳幼児221名のデータを得た。男女の比率はほぼ等しく収集されている。年齢については、得点の変化が大きい0歳(0-6ヶ月、7-12ヶ月)で1歳や2歳の約2倍のデータが収集された。

項目分析

Table 2にSPI(0-6ヶ月用)の各項目の平均値と標準偏差、修正済み項目-合計相関、度数分布を示す。ほとんどの項目において平均値が3点より高い値を示している。一般の項目1(わたしの子どもは、一日中活

発である)や視覚の1(わたしの子どもは、キラキラした物を見るのを楽しむ)など、平均値が低い値を示す項目が散見されるが、これらの項目内容は、定型発達の乳幼児にも見られる反応であるためと考えられる。

修正済み項目-合計相関(項目得点と当該項目を除いた尺度得点の相関)は各項目が当該項目の属する尺度と一致した概念を測定しているかを評価するための指標であり、数値が高いほど尺度との適合性が高いことを意味する。ただし、SPIの0-6ヶ月用尺度では、セクションごとの検討は行われないため、ここでは象限と低閾値(感覚過敏象限と感覚回避象限の合計)との相関のみ算出した。全般的に.20以上の値を示す項目が多いが、一部に.20を下回る項目や負の値を示す項目が見られる。ただし、0-6ヶ月については、50名のデータしか収集されていないため、これらの値の低さが項目の適合の悪さによるものか、サンプルサイズの不足による相関の不安定性によるものかについては、さらに多くのデータを収集して検討する必要がある。

度数分布を見ると、一部の項目で欠損の割合が多いことが見て取れる。特に、一般2、一般6、前庭覚1については、全体の10%にあたる5名が欠損となっている。これらの項目はいずれも、他の子どもと比べたときの反応の程度を尋ねる項目であるため、0-6ヶ月児の親で他の子どもの様子を知らない場合には、これらの項目への回答が難しかったものと考えられる。

続いて、Table 3にSPI(7-36ヶ月用)の各項目の平均値と標準偏差、修正済み項目-合計相関、度数分布を示す。ほとんどの項目で平均値が4点という高い値を示してい

る。これは、0-6ヶ月と比較して発達が進んだことで、より子どもの反応がはっきりし、評定が安定したためであると考えられる。前庭覚の項目2など、平均値が低い項目も見られるが、0-6ヶ月用と同様、定型発達の子どもにも広く見られる反応であるためと考えられる。

項目-合計相関については、セクション単位、低閾値を含む象限単位のいずれにおいても全般的に十分な値が示された。しかし、一般1および聴覚8については、象限との相関で.10前後の低い値が示されている。一般1は「予定が変わると、わたしの子どもの行動は悪化する」という項目であり、感覚過敏の他の項目とは異なり、感覚レベルの反応を問う項目ではないため、相関が低くなったものと考えられる。聴覚8は、感覚回避象限で唯一聴覚セクションの項目であるため、相関が低くなった可能性がある。

欠損の割合を見ると、一部の項目でやや高くなっている。共通点として、鏡、テレビ、コップなど、幼い乳児ではまだ接触の機会が少ない事物に関する記述を含んでいる項目や、「座る」「歩いたり這ったりする」など発達の個人差によっては7ヶ月時点でまだ獲得されていない可能性のある行動が含まれている項目で欠損の割合が高くなっている。

内的整合性

Table 4に、各象限の月齢帯ごとの α 係数を示す。0-6ヶ月については、4象限のいずれも.60を下回る値を示しており、特に感覚回避は.35と低い値を示している。上述の項目分析において、感覚回避を構成する項目は平均値が上限の5に近いものが多く、天

井効果によって得点の分散が小さくなったことが α 係数の低さにつながっていると考えられる。他の象限についても、50名のサンプルから得られた値であるため、今後より多くのデータを収集し、安定的な結果を得る必要がある。7-36ヶ月については、一部低い値が見られるが、全体ではいずれの象限も.60を上回っており、許容しうる範囲にある。

Table 5に各セクションの月齢帯ごとの α 係数を示す(0-6ヶ月はセクションごとのスコアリングをしないため α 係数を算出してない)。月齢7-12ヶ月の口腔感覚、13-24ヶ月の一般が低い値を示しているが、年齢帯を合併した値ではいずれも.50を超える値を示している。象限よりも尺度ごとの項目数が少ないこともあり、係数は全体に低い値が示されている。

得点分布

Figure 1にSPI(0-6ヶ月)の各象限の尺度得点のヒストグラムを示す。Figure 2にSPI(7-36ヶ月)の各象限の尺度得点のヒストグラム、Figure 3に各セクションの尺度得点のヒストグラムを示す。なお、各月齢帯のサンプルサイズが小さかったため、年齢帯はプールした状態でヒストグラムを作成した。一般群サンプルに実施したため、全体的な傾向として、ピークが右に偏った分布となっている(感覚探究を除く)。

尺度得点の標準値

Table 6に月齢帯ごとの各象限の尺度得点の平均値・標準偏差を示す。0-6ヶ月と7-36ヶ月では項目構成が異なるため直接の比較はできない。7-36ヶ月における変化を見る

と、いずれの象限も月齢が上がるにつれて得点が高くなっているが、特に感覚探究では7-12ヶ月から25-36ヶ月にかけて約1.5SD分の変化が見られ、発達の影響を受けやすい象限であることが見て取れる。他の象限は月齢による変化が0.5SD程度までに留まっており、発達の影響は限定的であると言える。

Table 7にセクションごとの尺度得点の平均値・標準偏差を示す。いずれのセクションにおいても、月齢が上がるにつれ、得点が高くなる傾向が見て取れる(一般は逆)。象限よりも発達の変化が大きく、いずれのセクションも7-12ヶ月から25-36ヶ月にかけて1SD前後の変化が見られる。

2. SP(児童版)

1) SPの標準化

サンプル

Table 8に10歳以下の一般群561名、Table 9に11歳以上の一般群880名の参加者の内訳を示す。10歳以下については、性別、年齢ともにほぼ等比率で収集されている。11歳以上については、82歳までの幅広い年齢のデータを収集した。

項目分析

原版におけるSPの適用年齢は10歳までであるため、項目分析には10歳以下のデータのみ用いた。Table 10に各項目の記述平均、標準偏差、項目-合計相関、度数分布を示す。ほとんどの項目において、4点以上と高い平均値であった。項目-合計相関、セクション単位で一部にやや低い値が見られたが、因子単位ではいずれの項目も十分な値が示された。

度数分布を見ると、ほとんどの項目は欠損の割合が 5% (28) 以下となっているが、感情・社会 1、感情・社会 2、行為 2 の 3 項目は 5% を超える欠損が見られる。これらの項目は、「自分があまり好きになれないようだ」「大人」になれない」など幼い子どもについては評定の難しい内容や「書いたものが読みづらい」という幼児では機会の少ない内容が含まれるために欠損が多くなったものと考えられる。

得点分布

SP (10 歳以下) のセクションごとのヒストグラムを Figure 4 に、因子ごとのヒストグラムを Figure 5 に示す。SP (11 歳以上) のセクションごとのヒストグラムを Figure 6 に、因子ごとのヒストグラムを Figure 7 に示す。一般群サンプルに実施したため、全ての下位尺度において、ピークが右に偏った分布となっている。

内的整合性

Table 11 に SP (10 歳以下) の年齢帯ごとの各セクションの α 係数を示す。いずれの年齢帯・セクションにおいても許容しうるレベルの値か高い値が得られ、内的整合性の高さが示された。ただし、5:0-6:11 歳の閾値のみ、低い値を示している。この理由としては、閾値を構成する項目が 3 項目しかなく、さらに、閾値の項目 2、3 が著しい天井効果を示しているためであると考えられるが、年齢帯を合併した場合には、60 前後の値を示しており、サンプルの影響も大きいと考えられる。続いて、Table 12 に SP (10 歳以下) の年齢帯ごとの各因子の α 係数を示す。いずれの年齢帯・因子についても高

い値が得られ、内的整合性の高さが示された。

次に、SP (11 歳以上) のデータを用いた内的整合性の検討についての結果を報告する。Table 13 に年齢帯ごとの各セクション α 係数を示す。ほとんどの年齢帯・因子において、高い値が示され、尺度の内的整合性が確認された。SP (10 歳以下) と同様、閾値の値がやや低めであるが、全体として見ると許容しうるレベルであると考えられる。続いて Table 14 に年齢帯ごとの各因子の α 係数を示した。いずれの年齢帯・因子についても許容しうるレベルか高い値が得られ、内的整合性の高さが示された。

尺度得点の標準値

まず、SP (10 歳以下) の結果について報告する。Table 15 に年齢帯ごとの各セクションの平均値と標準偏差、Table 16 に年齢帯ごとの各因子の平均値と標準偏差を示す。平均値の推移を見ると、一部の尺度は発達に伴って尺度得点に変化しており、その変化幅は感覚探究のみ 1SD 程度と大きい、その他は最大で 0.4SD 程度までに留まっており、発達の影響は全体に小さいと言える。

2) SP の妥当性 サンプル

Table 17 に臨床群の参加者の内訳を示す。一般群との比較においては、知的障害の診断の有無と、ASD の診断の有無により、MR 群 (知的障害あり・非 ASD ; 32 名)、高機能 ASD 群 (知的障害なし・ASD ; 79 名)、低機能 ASD 群 (知的障害あり・ASD ; 79 名) との対比を行った。知的障害なし・非 ASD 群については、人数が非常に少なかった。

たため、一般群との比較は行わなかった。知的水準、自閉的特性、適応行動との関連に関する分析では全てのサンプル(312名)を合併して分析を行った。

診断ごとの得点分布

Figure 8–Figure 13 に、臨床群ごとに各セクション・各因子の尺度得点のヒストグラムを示した。サンプルサイズが小さいものの、一般群のヒストグラムと比較して、左裾の厚さが増し、低得点帯に位置するデータが多くなっていることがわかる。特に ASD 群（高機能・低機能）は、MR 群よりもその傾向が顕著である。

診断ごとの尺度得点

一般群と各臨床群との尺度得点の比較を行った。その際、検定に必要となるサンプルサイズを確保するために、年齢帯をプールした上で分析を行った。まず、Table 18、Table 19 に MR 群と一般群との尺度得点についての t 検定の結果を、効果量 d とともに示した。23 個ある下位尺度（14 セクション、9 因子）のうち、口腔感覚過敏、寡動、前庭感覚、活動レベルの 4 尺度を除く 19 の下位尺度において、MR 群の方が一般群よりも尺度得点が有意に低かった。効果量 d は二群間の平均値の差を一般群の標準偏差で除した指標で、一般に 0.2～.05 で小さい差、0.5～0.8 程度で中程度の差、0.8 以上で大きな差を示すとされる。セクション単位では、 d が 0.2～0.5 の小さな差を示すものが 5 セクション、0.5～0.8 の中程度の差を示すものが 3 セクション、0.8 以上の大きな差を示すものが 6 セクションであった。因子単位では d が 0.2～0.5 の値を示すものが

2 因子、0.5～0.8 の値を示すものが 2 因子、0.8 以上の値を示すものが 4 因子であった。

次に、Table 20、Table 21 に高機能 ASD 群、と一般群との比較の結果を示した。全ての下位尺度において、高機能 ASD 群の方が一般群よりも有意に得点が低かった。効果量 d の値を見ると、セクション単位では 0.2～0.5 の小さい差を示すものが 2 セクション、0.5～0.8 の中程度の差を示すものが 3 セクション、0.8 以上の大きい差を示すものが 9 セクションであり、全般的に MR 群より顕著な差が見られた。因子単位でも、0.2～0.5 の小さい差が 2 因子、0.5～0.8 の中程度の差が 1 因子、0.8 以上の大きい差が 6 因子であり、MR 群より顕著な差を示した。

Table 22、Table 23 に低機能 ASD 群と一般群との比較の結果を示した。全ての下位尺度において、低機能 ASD 群の方が一般群よりも有意に得点が低かった。効果量 d の値を見ると、セクション単位、因子単位のいずれにおいても、全てが 0.8 以上の大きい差を示しており、MR 群や高機能 ASD 群よりも顕著な差が示された。

知的水準・自閉的特性との関連

臨床群における SP の各セクション・因子と、知能指数および PARS との関連を検討した (Table 24、Table 25)。ここでは臨床群の分類をプールし、臨床群全体の傾向を分析した。その結果、IQ との相関は、部分的に有意であったが、係数は低いまたは中程度の値に留まった。一方、PARS との相関については、全般的に有意な相関が見られ、特に、敏感性尺度との相関が中程度から高い値を示していた。また、PARS 下位尺度に含まれる全項目の合計と、SP との間には、

全体的に中程度の相関が得られた。

適応行動・不適応行動との関連

IQ、PARSと同様に、VABS-IIによって測定された適応行動・不適応行動との関連を検討した (Table 26、Table 27)。適応行動との間には、顕著な相関関係が示されなかったが、コミュニケーション領域や社会性領域は、部分的に SP と弱い相関を示した。一方、不適応行動尺度とは全般的に中程度から強い相関が示された。

3. SPA (青年成人版)

1) SPA の標準化

サンプル

一般群の参加者 1226 名の内訳を Table 28 に示す。やや女性の比率が高いが、年齢は 11 歳から 82 歳まで幅広くサンプリングした。

項目分析

Table 29 に各項目の記述平均、標準偏差、項目-合計相関、度数分布を示す。全ての項目が 5 件法の中点である 3 点を下回っており、やや低い平均値を示していた。また、象限ごとの項目-合計相関も全ての項目で .20 を上回る値が示されており、著しく適合の低い項目は見られなかった。欠損の割合も、全ての項目が 5% を下回っており、特に顕著に欠損の見られる項目はなかった。

内的整合性

Table 30 に、年齢帯ごとの各象限の α 係数を示す。いずれの年齢帯・象限においても、 α 係数は .80 以上の高い値を示しており、尺度の内的整合性が確認された。

得点分布

各象限の尺度得点のヒストグラムを Figure 14 に示す。全体に正規分布に近い形を示している。

尺度得点の標準値

Table 31 に、各年齢帯における象限ごとの尺度得点の平均値と標準偏差を示す。いずれの象限も年齢による変化は小さく、変化幅は $0.3SD$ 程度までに留まっている。

2) SPA の妥当性

サンプル

臨床群のサンプルを Table 32 に示す。SP と同様、一般群との比較においては、知的障害の診断の有無と、ASD の診断の有無により、MR 群 (知的障害あり・非 ASD ; 54 名)、高機能 ASD 群 (知的障害なし・ASD ; 27 名)、低機能 ASD 群 (知的障害あり・ASD ; 7 名) との対比を行った。知的水準、自閉的特性、適応行動との関連に関する分析では全てのサンプル (178 名) を合併して分析を行った。

診断ごとの得点分布

Figure 15–Figure 17 に、臨床群ごとの各象限の尺度得点のヒストグラムを示した。

診断ごとの尺度得点

一般群と各臨床群との、各象限の尺度得点の比較を行った。SP と同様、サンプルサイズの確保のために、年齢帯をプールした上で分析した。まず、Table 33 に MR 群と一般群との尺度得点についての t 検定の結果を、効果量 d とともに示した。その結果、

いずれの象限についても、両群の平均値に有意な差は見られなかった。

次に、Table 34 に高機能 ASD 群との比較結果を示した。全ての象限において、一般群との間に有意な差が見られ、感覚探究は一般群が高機能 ASD 群より高く、他の 3 象限は高機能 ASD 群が一般群より高かった。効果量 d は 0.35–0.78 と、小さいまたは中程度の差を示している。最後に、Table 35 に低機能 ASD 群と一般群との比較の結果を示した。低登録象限においてのみ、一般群との有意な平均値差が見られ、低機能 ASD 群が一般群より高かった。効果量 d の値は、0.61 と中程度であった。

知的水準・自閉的特性との関連

臨床群における SPA の各象限と、知能指数および PARS との関連を検討した (Table 36)。ここでは臨床群の分類をプールし、臨床群全体の傾向を分析した。その結果、各象限と IQ には有意な相関が見られなかった。同様に、PARS の各因子および合計点とも有意な相関は見られなかった。

適応行動・不適応行動との関連

IQ、PARS と同様に、適応行動・不適応行動との関連を検討した (Table 37)。適応行動との間には、低登録が運動との間に有意ではあるが弱い相関を示しただけであった。一方、不適応行動尺度とは感覚探究を除く 3 象限が中程度の相関を示した。

D. 考察

1. SPI (乳幼児版) について

0-36 ヶ月の乳幼児 221 名から得たデータをもとに、SPI の標準化を行った。

項目分析の結果、大部分の項目は尺度と十分な適合 (項目-合計相関) を示したが、一部の項目でやや適合が低かった。特に 0-6 ヶ月用項目では、低い適合を示す項目の数が多かったが、0-6 ヶ月は 50 名のデータしか得られていないため、この結果が項目の適合の悪さを意味するものか、サンプルサイズの不足による結果の不安定性によるものかについては、さらなるデータ収集を行って検討する必要がある。また、幼い乳児では接触の少ない事物や獲得されていない行動に関する記述を含んだ項目で欠損が多い傾向が見られた。こうした合理的な理由による欠損をどのように処理すべきか、今後の検討の余地が残されている。

象限ごとの内的整合性は、7-36 ヶ月については全体に許容しうる範囲の結果が得られたが、0-6 ヶ月については全体に α 係数の値が低く、感覚回避では .35 という低い値が示された。この点についても、データの追加収集によって更なる検討が必要である。一方、セクションごとの内的整合性 (7-36 ヶ月のみ) は、象限ごとの内的整合性より全体に値が低かった。セクションは象限よりも区分が細かく、全体に項目数が少ないことが一因と考えられる。また、セクションは感覚モダリティ (視覚、聴覚、触覚など) に基づく区分であり、反応の心理行動学的性質に基づく区分ではなく、概念的な一貫性が低いため、内的整合性が低くなりやすいことも原因として考えられる。

象限・セクションごとの尺度得点の分布を見ると、正規分布に近い形を示すものが多かったが、一般群サンプルから得られたデータであるため、全体的にやや右に偏った分布となっていた。

月齢帯ごとの尺度得点の平均値は、いずれの象限も7-36ヶ月にかけて月齢が上がるほど得点が高くなっていった。特に感覚探究は7-12ヶ月から25-36ヶ月にかけて顕著な変化が見られ、発達の影響を受けやすいことが示唆された。しかし、他の象限は変化幅が0.5SD程度までに留まり、発達の影響は限定的であることが示された。

SPIについては、臨床群のデータが十分に得られず、妥当性の検証を行うことができなかった。SPIの適用範囲である0-2歳の時期にはまだ発達障害などの医学的診断を受けている子どもが少ないため、この時期の臨床群データを収集することは容易でない。しかし、一般群を対象に発達障害傾向に関するアセスメントや感覚の問題に関する他の尺度を実施してSPIとの関連を検討するなどの方法で妥当性を検証することは可能であり、今後の課題として検討を続けていく必要がある。

2. SP（児童版）について

3-10歳の561名と11-82歳の880名から得た一般群データと312名の臨床群データをもとに、SPIの標準化および妥当性の検証を行った。

項目分析の結果、セクション単位では尺度との適合（項目-合計相関）がやや低い項目が一部に見られたが、因子分析結果に基づく9因子区分では、いずれの項目も十分な適合を示した。欠損の割合も全体に低かったが、一部、幼い幼児では評定が難しいと思われる内容の項目でやや欠損が多い傾向が見られた。SPIと同様、こうした系統的な欠損への対処法を検討する必要がある。

内的整合性は、10歳以下、11歳以上のい

ずれもセクション単位では一部に低い値を示す尺度が見られたが、因子単位ではいずれの尺度でもほぼ十分な値が得られた。原版では10歳までがSPの適用年齢となっているが、この結果から、11歳以上に適用した場合にも同様の信頼性をもって測定を行えることが示された。

尺度得点の分布は、いずれの下位尺度も、ピークが右に偏った分布となっており、11歳以上では特にその傾向が顕著であった。ただし、SPはSPIやSPAと異なり、得点が低い場合にのみ異常判定の対象となるため、このような分布の形状であっても、判定に特別な支障をきたすことはない。

尺度得点の平均値は、全体に年齢が上がるほど得点が増える傾向が見られたが、その変化幅は感覚探究で1SD程度の変化が見られる以外はいずれの尺度も0.4SD程度までに留まっており、発達の影響は小さいと言える。SPでは年齢ごとの平均値とSDに基づいてカットオフ値が設けられるが、一般に各年齢帯50~100名程度のサンプルでは多くのランダム誤差が混入するため、サンプルの平均値やSDをそのまま母集団の推定値と見なすことはできない。本研究においても年齢帯ごとの平均値やSDには不自然な逆転や凹凸が見られる。こうした点については、locally estimated scatterplot smoothingなどの方法による平滑化を行うことで、より妥当な推定値を得る必要がある。

臨床群（MR群、高機能ASD群、低機能ASD群）と一般群の尺度得点を比較したところ、全般的にMR群よりも高機能ASD群、高機能ASD群よりも低機能ASD群で、一般群との顕著な差が見られた。また、IQお