

目 次

はじめに	2
1) 本研究の位置づけ	2
2) 本研究の目的	3
3) 本論文の構成	5
TKW 式検査の下位検査項目とその結果，および概念的妥当性の検証	6
1. TKW 式検査の下位検査項目，施行対象，および結果	6
2. TKW 式検査の均質性についての因子分析による検証	7
3. TKW 式検査が作動記憶理論に立脚していることを確認するための分析	8
1) 最大作動記憶デマンド数による理論的項目困難度との適合度の検証	8
2) Guttman の尺度化理論との適合度の検証	9
3) 項目反応理論の適用による推定尺度値との適合度の検証	9
4. 病型の分類	10
TKW 式検査と既存の評価法との比較	10
1) 対象と施行手続き	10
2) 結果	11
3) TKW 式検査と既存の評価法との比較による併存的妥当性の検証	12
4) 痴呆の重症度診断テストとしての予測的妥当性と信頼性の検証	12
臨床場面における実用化	13
1) コンサイス版	13
2) シングル版	15
3) 既存の評価法とコンサイス版，およびシングル版との比較	16
あとがき	17
 図表	
本論文の目次	19
研究業績一覧	22
引用文献	23

はじめに

1) 本研究の位置づけ

筆者の指導教授である並木博先生は、1991 年春、前任校慶應義塾大学において、同大学医学部出身の東海大学神経内科の教授たちと新しい痴呆診断のための神経心理学的検査の開発を目的とする共同研究のプロジェクトに着手した。本研究は、このプロジェクトの一環として行なわれたものである。本プロジェクトには、これまで幾人もの方々が、学部卒業論文、あるいは修士論文の実験として参加してきた。筆者は、1995 年に本学大学院教育学研究科に入学して以来、修士課程、博士課程を通じ、一貫してこのプロジェクトに参加し、検査データの収集と検査構成のための分析に従事してきた。本論文は、筆者のこのプロジェクトにおける過去 9 年間の研究成果をまとめたものである。

本プロジェクトは、神経内科側からの以下のような要請に答えるべく立案された。すなわち、現在わが国の医療現場で広く用いられている改訂長谷川式簡易知能評価スケール (HDS-R) は、実施が容易である点を大きな特徴としているが、測定対象が記憶能力に偏りすぎる傾向があるため、より広く認知能力を捉え得る検査の開発が必要である。心理学の立場から、HDS-R の問題点をあげれば、記憶能力への偏りとともに、以下の 2 点が指摘できる。すなわち、検査を構成する項目はいずれも従来の知能検査の中から、特に記憶関連の項目を選定したものであり、理論的背景を欠いていること、そして、検査結果の得点化は、各項目より得られる素点に適宜重みづけをして加算した合計得点によっているが、この手続きに心理測定学的理論を欠いていることである。

また、医療現場で用いられている行動評価法は、記憶、見当識、判断力と問題解決、社会適応、家族状況と趣味・関心、および介護状況について、家族や介護者などから得た情報を組み合わせることにより評価するものであり、その方法の信頼性、妥当性ととともに疑問視せざるを得ない。さらに、検査法の中には、認知能力その他を多角的に捉えようとするあまり、被検査者に多大の負担を強いるものも見られる。

以上のような、現在医療現場で使用されている検査法のいくつかの問題点を解消し、より良い診断検査を開発することが本プロジェクトの目標であった。

検査項目の作成、選択における理論的基礎としては、R. Case の提唱した作動記憶理論に依った Case は Piaget の認知発達理論の基本概念である「構造」概念を操作化するために、作動記憶容量 (working memory, 以下 WM とする) という構成概念を導入した。このような Piaget を超える研究実績が評価を得、Case は新ピアジェ派の代表格とされてきた。さら

に、Case は、WM に基づく発達的研究の成果を補償的教授方法の開発に応用することを試みて、良い結果を得ている。つまり、Case は、発達心理学者でありながら、教育心理学者でもあった。そして、このような実践的研究の結果がまた逆に、理論モデルの改善にも繋がったと考えられる。

本プロジェクトは、もともと認知発達心理学的研究の中での理論化を生涯発達心理学の枠組みの中にまで延長し、適用する試みである。また、新しい検査項目の作成にあたり、並木らによるこれまでの認知発達心理学的研究（並木，1964；並木ほか，1976）の積み重ねがあった。その中で用いられた実験課題が痴呆と WM を関連づける仮説の下に、痴呆診断に役立ち得るという理論的見通しと、またそれを裏づける Baddeley らの実験的研究を踏まえて、本プロジェクトは進められてきた。

また一方では、WM 理論は、従来の一般知能にかわって、学習困難児（LD）や注意欠陥・多動性障害児（AD/HD）の行動にかかわる説明概念として注目されており、筆者がかつて考案した（米倉，2001）ようにこの理論の汎用性は特筆されるべきである。

次に、検査得点の尺度化にあたり、本検査は当初より項目反応理論の適用をもくろんでいたが、結果的にこれが可能となり、この理論の特徴を生かして、医療現場における実用性を高めた検査の作成に成功した。WM 理論も項目反応理論も心理学の中で展開してきたものであるが、これらの理論に基づく本検査法が医学会でその効用を認められるならば、本プロジェクトは、その目的を達成したことになる。

最後に、教育心理学の枠組みの中に本研究を位置づけるならば、教育心理学に含まれる認知発達、思考、評価・測定、臨床の下位領域にかかわると言える。本検査が痴呆の重症度や病型の診断に用いられるならば、評価・測定と臨床にかかわり、痴呆患者のリハビリテーション・プログラムの効果の測定に用いられるならば、これは教育評価にかかわる。また、痴呆の病態のメカニズムの解明を目指すならば、認知発達、思考、および臨床の領域にかかわる。このように、本検査の開発は、教育心理学の諸領域の研究成果の医学への応用として位置づけることができる。そして、そこで得られた結果が逆に教育心理学の理論の修正と発展に資するものと考えられる。

2) 本研究の目的

本研究の目的は、大きく分けて3つある。第1の目的は、現在用いられている痴呆の評価法の問題点の改善を目指した検査法を作成することである。筆者は、WM容量という唯一

の心理学的概念を測る「ものさし」を用いて、痴呆の重症度の診断を試みる。この「ものさし」は、患者集団の所在地が、例えば東京と沖縄というように生活習慣等が大きく異なっているとしても使えるものでなければならない。これまで、痴呆症状、特にアルツハイマー型痴呆の症状にWMが関与しているという実験心理学的研究がBaddeley ほか(1991)によって報告されている。しかし、Baddeleyの提唱するWMモデルは、複数のコンポーネントからなるので、脳の機能を全体的に捉えるには不便である。そこで、「中央のWMが1つあり、容量に限界があり、入力した情報を処理したり処理結果を貯蔵したりするために使われる融通性のあるwork space」という点でBaddeleyと同様のモデルを提唱し、WMを量的に捉えたCaseのモデルを用いることにする。Case(1980)は、「ある検査課題の遂行が可能であるのは、被検査者のWM容量が課題の最大デマンドと等しいか、あるいはより大きい場合のみである」としたが、この理論は、Guttman(1900,1947)による尺度分析法(scalogram analysis, 以下SAとする)に一致している。もし、筆者らの開発した神経心理学的検査のこのSAに対する適合の良さが確認できれば、1つの心理学的次元を測定対象としていることが実験データにより裏づけされることになる。本研究では、WM理論と項目反応理論の適用により、概念的妥当性の高い痴呆の重症度診断検査(Tokai-Keio-Waseda式検査、並木ほか、2002、以下、TKW式検査とする)を開発することを目的の一つとした。

また、Caseは、WM容量が貯蔵スペースと情報処理スペースからなるとし、WM容量は2歳以降、個々に一定となるが、課題処理操作の効率が高くなり自動化されれば、その実質的な容量が増えるとした。本研究の第2の目的は、TKW式検査が、WMという唯一の心理学的次元を捉えていることを踏まえ、この下位検査項目を訓練課題として用い、課題に対する自動化を促し、WMの実質的な容量の増加により、痴呆症状の改善を目指すことである。つまり、痴呆患者に対して具体的な処方を行ない、その適切さを検証することが第2の目的である。

そして、第3の目的は、臨床場面におけるTKW式検査の実用化である。TKW式検査は、WM理論と項目反応理論に裏づけされており、項目単位で扱うことが可能となった。そこで、現在臨床現場で用いられているテスト法の問題点を改善し、痴呆の重症度に合わせることで、短くて情報量の多いtailored test、および痴呆患者の中で最も多いアルツハイマー型痴呆と脳血管性痴呆との分類を目的とした簡易検査法の作成を試みる。

3) 本論文の構成

本論文は、以下のように構成されている。

第2章では、痴呆の病態を神経心理学的、および認知心理学的に捉える。そして、現在用いられている代表的な神経心理学的検査の特徴を記し、本研究で用いた既存の評価法について筆者が感じた問題点を述べる。その上で、Lezak (1982, 1995) の提唱する「遂行機能」が痴呆という病態を捉える概念として妥当であることを提案する。

第3章では、WM理論の時代的な変遷についてLogie (1996) を引用し、次に、筆者らが掘りどころとしたBaddeleyとCaseによるWMモデルを説明する。さらに、知能因子構造論について説明し、一般知能因子Gf がWMと密接な関係があるというEngle (1999) ほかの報告を紹介する。

第4章では、本研究で用いたGuttmanの尺度分析法と、項目反応理論の1つである2パラメータ・ロジスティック・モデルについてGreen (1956)、池田 (1994) ほかを引用する。

第5章では、並木ほかが開発した神経心理学的検査がTKW式検査として構成されるまでの経緯を説明する。

第6章で、痴呆症状を主訴とする患者196名に、TKW式検査を施行した結果からこの検査の概念的妥当性を検証する。

第7章で、痴呆の重症度診断の外部基準として既存の評価法を用い、TKW式検査の併存的妥当性を確認する。そして、発症メカニズムの異なるDATとVDの病型の識別力の程度を分析する。

第8章で、Caseの理論に基づき、操作に慣れさせ、効率化を促す訓練によるWMの実質的容量の増加の可能性について検証する。訓練に用いる課題として、TKW式検査の下位検査項目のうちからWMに大きな負荷をかけるものを選び、さらにその中からDATの反応パターンの特徴をとらえた「事物の階層構造的分類」と「漢字の仲間はずれの選択」を選定した。この課題を拡張した訓練課題による痴呆患者への訓練効果を報告する。

第9章では、臨床現場における更なる実用化を目指して作成した、重症度別、および病型別の簡易検査法（コンサイス版とシングル版）の有用性を検証する。

第10章で、まとめ、および今後の展開について述べる。

本概要書においては、以下に、 で TKW 式検査を構成する下位検査項目、施行結果、および概念的妥当性の検証、 で TKW 式検査と既存の評価法との比較、 で臨床場面にお

ける実用化，そして であとがきについてその概要を述べ，最後に，本論文の目次，図表目次，研究業績，そして引用文献を付した．

TKW 式検査の下位検査項目とその結果，および概念的妥当性の検証

1. TKW 式検査の下位検査項目，施行対象，および結果

検査項目：研究 ， の結果に基づき，各項目の困難度，被験者への指示の伝達の難しさ，研究 から得られた因子構造，研究 ， の比較のための多変量分散分析と判別分析の結果，WM 負荷の大きい 4 コの検査項目に数唱を加えた 5 コの検査項目により TKW 式検査を構成した．

表 1 に示すように，5 コの検査項目は， 事物の名称と階層構造的分類， おはじきの計数と数の再現， 漢字の読みと仲間外れ， 迷路， 数唱である．これらの検査項目より， 事物の名称， 事物の階層構造的分類， おはじきの計数， おはじきの数の再現， 漢字の読み， 漢字の仲間外れ， 迷路， 数唱，以上合計 8 コの検査項目得点が得られる．そして，8 コの検査項目得点は，12 コの名称，4 コの階層構造分類，2 コの計数，2 コの再現，10 コの漢字読み，2 コの漢字の仲間はずれ，3 コの迷路，および 1 コの順唱の計 36 コの下位検査項目得点から構成される．さらに，4 コの階層構造分類と 2 コの漢字の仲間はずれの選択は，それぞれ分類あるいは選択，分類・選択の理由付け，分類・選択に要した時間からなり，3 コの迷路は，迷路と所要時間からなる．つまり，8 コの検査項目得点は，12 コの名称，階層構造分類が $4 \times 3 = 12$ ，2 コの計数，2 コの再現，10 コの漢字読み，漢字の仲間はずれの選択が 2×3 ，迷路が 3×2 ，そして 1 コの順唱の計 51 コの下位検査項目得点から構成される．

対象：1995 年 4 月から 2003 年 8 月までに，某大学医学部付属病院神経内科および都内某病院内科を受診した物忘れを主訴とし，痴呆が疑われる患者 196 名（平均 73 歳，29-93 歳）．うち男性は 71 名（平均 70 歳，29-88 歳），女性は 125 名（平均 75 歳，46-93 歳）である．臨床的および画像診断上 DAT と診断されたのは 85 名（平均 75 歳，29-92 歳），VD と診断されたのは 36 名（平均 74 歳，46-90 歳），合計 121 名（平均 75 歳，29-92 歳）であった．ほかに，年齢相応が 12 名（平均 66 歳，37-90 歳）と診断名不明が 62 名（平均 72 歳，57-93 歳）である．

結果：全検査の所要時間は 18.6 ± 12.3 分（8～60 分）であった。診断名別では、DAT 群で 21.8 ± 13.6 分（9～60 分）、VD 群で 16.5 ± 13.9 分（8～50 分）、年齢相応群で 12.8 ± 2.0 分（10～17 分）と、DAT 群で長かった。全検査所要時間の平均値における DAT 群と VD 群間の差は 5% 水準で有意であった。

素点合計得点の平均値と標準偏差、および幅は、全体で 45.6 ± 17 （4～66）、DAT 群で 39.9 ± 16.1 （4～66）、VD 群で 48.2 ± 18.1 （14～66）、年齢相応群で 62.1 ± 3.2 （55～66）であった。素点合計得点の平均値が DAT 群と VD 群間で 0.5% 水準で有意差があったのは、全検査合計得点（36 項目）、事物の階層構造的分類の理由づけ、漢字の仲間はずれの理由づけ、そして迷路 1 であった。素点合計得点の平均値が DAT 群と年齢相応群間で 0.5% 水準で有意差があったのは、全検査合計得点、事物の名称と階層構造的分類、およびその理由づけ、おはじきの数の再現、漢字の仲間はずれとその理由づけ、そして迷路とその所要時間であった。

Cronbach's α は、0.91 であった。再検査信頼性係数は、全検査素点合計得点間の相関係数が $r = 0.88$ ($p < .0001$)、下位検査項目得点間の相関係数が、 $r = 0.19$ であった「おはじきの数え」を除き $r = 0.65$ ($p < .0001$) ～ $r = 0.86$ ($p < .0001$) であった。以上のように、信頼性は、十分高いことが確認された。

2. TKW 式検査の均質性についての因子分析による検証

8 コの検査項目得点、および 13 コの下位検査項目得点の因子分析を主因子法で行った。表 2 に示すように、検査項目が 8 コの場合は、第一因子の固有値あるいは因子寄与が 4.26 と大きく、第二因子は 0.87 となり、その後急速に 0 に近づくので、Kaiser の基準（1960）に従い第二因子以下は無視することができる。つまり、これらの検査項目得点が唯一の共通次元を測っているということになる。また、これまで被検査者のデータ数の増加とともに、40 名、80 名、100 名につき因子分析を繰り返し行った結果、因子パターンは同じで、均質テストの因子構造は十分安定したものと考えられた。なお、Cronbach's α は 0.86 と、内的整合性も高かった。

情報処理内容の異なる 13 コの下位検査項目得点の因子分析の結果は、固有値が第 1 因子で 6.46、第 2 因子で 1.25 であった。このように、第 2 因子の因子寄与は Kaiser の基準の 1 以下とはならなかったが、第 1 因子負荷量がほぼ全ての項目で大きく、第 1 因子寄与率が 49.7% と大きいこと、8 コの検査項目得点では単一因子構造として安定していたこと、

Cronbach's α が 0.91 だったことなどからほぼ均質なテストとみなされる。そして、以下の理由でこの第一因子が WM の因子であると解釈されうる。即ち、因子負荷量の最も大きいものは、漢字の仲間はずれ、事物の階層構造分類、おはじきの数の再現、迷路と、いずれも WM に大きな負担をかける項目が並び、一方、因子負荷量の小さいものは、数唱、おはじきの計数、漢字の読みと WM の負担が少ない項目であったからである。

以上のように、TKW 式検査が WM という 1 つの心理学的次元を測る検査であることが確認できた。

3. TKW 式検査が作動記憶理論に立脚していることを確認するための分析

1) 最大作動記憶デマンド数による理論的項目困難度との適合度の検証

R. Case (1980) による WM 理論は、「中央の WM が 1 つあり、容量に限界があり、入力した情報を処理したり処理結果を貯蔵したりするために使われる融通性のある workspace」という点で Baddeley の WM モデルと共通している。異なるのは、処理能力容量の限界について、Baddeley らがコンポーネントの違いに焦点を当てたのに対し、Case は量化に焦点を当てたことである。

筆者は、痴呆の病態を、様々な症状の上位に位置する遂行機能の低下と考える。そして、「認知機能の一般的発達要因は、WM 容量の量化可能な水準であり、WM 容量の増加は、操作の自動化である」という Case の WM 理論に基づけば、痴呆症状を発達過程の終末段階における WM 容量の低下と解釈することができる。そこで、「Noelting のジュース問題を遂行するために必要な WM デマンド」に倣い、下位検査項目の理論的困難度を決定し、それと実測正答率との一致度により、痴呆のための TKW 式検査に対する WM 理論のあてはまりのよさを検証する。

下位検査項目得点の最大 WM デマンド数は、事物の名称が 5、事物の階層構造的分類 1 が 8、分類理由付けが 7、事物の階層構造的分類 2、3、4 が 7、おはじき計数 2 桁が 4 で 3 桁が 5、おはじき再現 2 桁が 6 で 3 桁が 7、漢字の読みが 4、仲間はずれの選択が 8 で選択の理由づけが 7、数唱 4 桁の場合が 6、迷路 1 が 7、迷路 2 が 9、迷路 3 が 10 となった(表 3)。

実測正答率による順位と下位検査項目数 13 コの WM 最大デマンド数による困難度の順位との順位相関係数は、 $r_s = 0.90$ ($p < .001$) と高く、項目困難度が WM 容量にかかる負荷の大きさによって説明可能であることが検証された。

2) Guttman の尺度化理論との適合度の検証

Case (1980) は、「ある検査課題の遂行が可能であるのは、被検査者の WM 容量が課題の最大デマンドと等しいか、あるいはより大きい場合のみである」としたが、この理論は、Guttman (1900, 1947) による尺度分析法 (scalogram analysis、以下、SA とする) に一致する。TKW 式検査の 13 コの下位検査項目得点の SA に対する再現性指数は $Rep_A = 0.931$, $Rep_B = 0.933$, $Rep_I = 0.840$ であった。この結果から、TKW 式検査が一次元尺度化可能であることと、Case の理論との適合の良さが検証された。あてはまりのよさの検証としては、この Guttman 尺度の再現性指数の方が、次の 2 パラメータ項目特性曲線への適合度検定²よりも厳しい。

3) 項目反応理論の適用による推定尺度値との適合度の検証

項目反応理論に用いた下位検査項目数 36 コの正答率による下位検査項目の困難度の順位と理論的困難度順位との順位相関係数は、 $r_s = 0.93$ ($p < .001$) と、WM 理論による困難度の当てはまりのよさが確認された (表 3)。

36 コの下位検査項目の項目特性曲線への適合度は、迷路 2 を除き、全体として良好であった。この結果に加え、ロジスティック関数を用いた項目特性曲線より厳しい Guttman 尺度への適合度が十分高かったことから、TKW 式検査が一次元尺度であることが検証された。

2 パラメータ・ロジスティック・モデルによる分析の結果、項目識別力の平均値と標準偏差は 1.18 ± 0.45 、項目困難度の平均値と標準偏差は -1.06 ± 0.36 だった。項目識別力および項目情報関数が最も高かったのは、事物の名称では「イヌ」、階層構造的分類 2, 3, 4 の中では動物を獣と鳥に分ける「分類 3」、漢字の読みの中では「桜」、仲間はずれの選択では「村」、「おはじきの再現 3 桁」、「迷路 2」であった。

項目反応理論の適用によって推定された項目困難度と正答率とが、最大 WM デマンド数によって決定された理論的困難度と対応すれば、TKW 式検査の概念的妥当性の高さが検証されることになる。結果は、推定された項目困難度と実測正答率との相関係数は、13 項目で $r = 0.98$ ($p < .001$)、36 項目で $r = 0.97$ ($p < .001$)、推定された項目困難度と最大 WM デマンド数によって決定された理論的困難度との順位相関係数は、13 項目の時 $r_s = 0.91$ ($p < .001$)、36 項目の時 $r_s = 0.95$ ($p < .001$) といずれも高く、TKW 式検査の概念的妥当性の高さが検証された。

図 1 に、実線でテスト特性曲線を、破線で標準誤差曲線を示した。困難度が -1.75 のと

きテスト情報関数は最大の 25.8 であった。標準誤差は、推定尺度値が - 3 から 0.98 まで 0.4 未満であった。

テスト信頼性係数は、0.93 であった。

4. 病型の分類

判別分析の結果

196 名のうち、臨床的および画像診断上 DAT と診断された 85 名(平均 75 歳, 29-92 歳), VD と診断された 36 名(平均 74 歳, 46-90 歳), 合計 121 名(平均 75 歳, 29-92 歳)について、判別分析を行った。

36 コの下位検査項目得点の素点合計得点による判別分析の結果、正しく分類されたのは、DAT85 名中 67 名 (78.8%), VD36 名中 32 名 (88.9%) の合計 99 名 (81.8%) だった。下位項目得点をさらに細かく分けて分析した結果、45 項目で 85.1% , 51 項目で 87.6% と項目数の増加に伴い、的中率が高くなった。51 項目による的中率は、DAT が 74 名の 87.1% , VD が 32 名の 88.9% であった。

病型の判別は、同じ推定尺度値における DAT 患者と VD 患者とでは、反応パターンが異なり、判別の的中率は、重症度と下位項目反応パターンの双方によって決定されることが確認された。下位項目反応パターンにおいては、事物の階層構造的分類と漢字の仲間はずれの選択が DAT と VD の病型分類に有効と考えられた。

TKW 式検査と既存の評価法との比較

TKW 式検査の痴呆の重症度診断テストとしての有用性を確認する為に、既存の行動評価法、およびテスト法とを比較し、所要時間、重症度の判定度、信頼性の点から検証する。

1) 対象と施行手続き

対象：1995 年 4 月から 2003 年 7 月までに、某大学医学部付属病院神経内科および都内某病院内科を受診したものの忘れを主訴とし、痴呆が疑われる患者 196 名のうち、既存の評価法を施行した被検査者数は、FAST が 117 名、CDR が 65 名、HDS-R が 136 名、MMSE が 74 名、ADAS-Jcog が 78 名、かなひろいテストが 69 名である。既存の評価法を施行した患者の平均年齢は、FAST が 72.7 歳 (29 ~ 91 歳), CDR が 69.8 歳 (29 ~ 90 歳), HDS-R が 73.3 歳 (29 ~ 92 歳), MMSE が 69.7 歳 (29 ~ 90 歳), ADAS-Jcog が 70.2 歳 (29 ~ 90 歳),

かなひろいテストが 69.5 歳 (29 ~ 90 歳) である .

施行手続き : 評価法のうち , テスト法については , 被検査者の負担を考慮し , TKW 式検査 , MMSE , HDS-R , かなひろいテスト , TKW 式検査 , MMSE , ADAS-Jcog と 2 つのセットに分けた . 1 つのセットを施行した後 , 被検査者の症状が変化しないと考えられる 2 , 3 週間以内に 2 セット目を施行した . セットの施行順序 , およびセット内のテストの施行順序はカウンターバランスをとった . テスト - 再テスト信頼性を確認するため , TKW 式検査と MMSE を と の両セットに入れた . 各テスト項目はそれぞれのマニュアルに沿い , 1 対 1 面接方式により施行した . FAST および CDR による評価は , 原則として , 医師が行った .

2) 結果

素点合計得点の平均値と標準偏差 (範囲) は , HDS-R が 16.8 ± 7.5 (3 ~ 30) , MMSE が 20.1 ± 6.4 (3 ~ 30) , ADAS-Jcog が 19.3 ± 10.9 (3 ~ 58.7) , かなひろいテストが 5.8 ± 8.9 (0 ~ 36) , FAST が 3.8 ± 1.4 (2 ~ 6) , CDR が 1.6 ± 0.8 (0.5 ~ 30) , そして TKW 式検査が 50.2 ± 17.7 (6 ~ 74) であった .

検査所要時間の平均値と標準偏差 (範囲) は , HDS-R が 5.4 分 ± 1.7 (2 ~ 11 分) , MMSE が 8 分 ± 2.3 (4 ~ 17 分) , ADAS-Jcog が 30 分 ± 8.3 (18 ~ 58 分) , かなひろいテストが 6 分 ± 1.5 (2 ~ 10 分) であった . 一方 , TKW 式検査は , 18.8 分 ± 12.3 (9-60) であった .

所要時間と得点との相関係数は , HDS-R が $r(59) = -0.483$ ($p < .001$) , MMSE が $r(69) = -0.336$ ($p < .001$) , ADAS-Jcog が $r(70) = 0.579$ ($p < .001$) といずれも成績のよい被検査者ほど所要時間が少ないという結果であった . 一方 , TKW 式検査は , 既存のテスト法とは異なり , 所要時間と素点合計得点との相関係数は , $r(196) = 0.376$ ($p < .001$) と , 成績が高いほど所要時間が長くなるという傾向を示した . この結果から , TKW 式検査が以下の点で他の検査と異なると考えられる .

時間制限をしない力量検査である .

結果が , 純粹に項目に対する反応となるように , 教示内容を完全に理解するまで説明し , 必要であれば , 練習させる .

手指の動きがスムーズでなかったり , 目が見えにくかったりと , 作業そのものに時間がかかる .

分類や仲間はずれの選択では、判断の理由を問うので、被検査者によっては「わからない」と言う代わりに、いろいろな言い訳をしたり、話題をそらしたりする。おはじきの再現では、間違えた場合、計数結果の数字を暗証するところから何度でもやり直す。

数唱課題は、2回連続して失敗するまで行う。したがって、能力の高い被検査者ほど多くの桁数を試行する。

迷路課題では被検査者の力量を見るため、検査者が間違いを指摘するのは、被検査者が間違いに気づかずにゴールに到達した場合のみとする。検査者は、被検査者が自分で間違いに気づくまで、あるいは give up するまで間違いを指摘せずに待つ。

スピードテストより力量テストの方が潜在能力をより多く発揮することができ (Rogosa, 1982), そして特に高齢者では、正確さを優先させるために、スピードを犠牲にする傾向が強い (McGrew, 1998) という理由からも、時間制限をしない TKW 式検査は痴呆の重症度診断検査として適している。

3) TKW 式検査と既存の評価法との比較による併存的妥当性の検証

TKW 式検査による推定尺度値、および素点合計得点を既存の評価法による得点と比較した。FAST、および CDR との相関係数は、5つのテスト方のうち TKW 式検査が最も高く、TKW 式検査が日常生活動作能力をよくとらえた検査法であることを示している。

既存のテスト法と TKW 式検査との相関係数は、「かなひろいテスト」を除き、 $r = 0.72$ ($p < .001$) から $r = 0.81$ ($p < .001$) と高く、併存的妥当性のあることが確認できた。

4) 痴呆の重症度診断テストとしての予測的妥当性と信頼性の検証

TKW 式検査による推定尺度値、および全検査項目素点合計得点と、FAST、および CDR の得点との対応を見ることにより、痴呆の重症度診断テストとしての妥当性を検証する。

TKW 式検査による推定尺度値、および素点合計得点の平均値は、FAST3 と 4、4 と 5 の間に 0.1% 水準、FAST2 と 3、5 と 6 の間に 0.5% 水準で有意差があった。また、CDR1 と 2、2 と 3 の間に 0.1% 水準、0.5 と 1 との間に 5~10% 水準で有意差があった。図 2 に TKW 全検査による推定尺度値と FAST、および CDR 得点との関係を示した。縦軸が FAST、および CDR 得点ごとの全検査による平均推定尺度値である。横軸に示した目盛りの上段が

FAST 得点，横軸の下段が CDR 得点である．FAST 得点別の全検査による平均推定尺度値を ，CDR 得点別の全検査による平均推定尺度値を ，そして ± 1 SD を直線で示した．平均推定尺度値は，疑い(CDR2)で 1.73，年齢相応(FAST2)で 1.53，軽度(CDR1，FAST3)で 0.94～1.29，中等度(FAST4，CDR2)で 0.20～0.26，やや高度(CDR2.5，FAST5)で -0.66～-0.49，高度または重度(FAST6，CDR3)で -1.39～-1.29 であった．このように，FAST と CDR による痴呆重症度の評価得点と推定尺度値との方向は一致しており，TKW 式検査が重症度診断検査として有効であることが確認できた．

TKW 全検査素点合計得点の再検査信頼性係数は， $r(56) = 0.884$ ($p < .0001$) であった．なお，MMSE の素点合計得点の再検査信頼性係数は， $r(37) = 0.934$ ($p < .0001$) であった．

臨床場面における実用化

TKW 式検査の検査項目が WM 容量という唯一の心理学的次元を測る検査であり，重症度診断として有用であることが検証された．しかし，力量テストである性質上，所要時間が約 18 分と長く，実用性にかけるという問題点がある．そこで，これまでの項目分析の結果を踏まえて，情報量をできるだけ多く保ちつつ，検査所要時間を短くし，被検査者の負担を軽減することにより，臨床現場での実用性を高めることを目的とした．

1) コンサイス版

項目の選定

コンサイス版の作成にあたり，第 1 の選定基準として，TKW 式検査において最も特徴的な均質性を損なわないよう，表 1 に示した 8 コの下位検査項目から下位検査項目得点を 1 つずつ選定した．第 2 の選定基準は，因子負荷量，項目情報関数，項目識別力，項目困難度，判別分析による構造行列の相関係数，の 5 つである．

から までの基準を選んだ理由は，

は，第 1 因子負荷量の大きい項目を選ぶことにより，テストの均質性を損なわないようにする．そして，第 1 因子負荷量の大きい項目ほど WM デマンドが大きく，従って，測定対象である WM 容量をより敏感に反映する．

は，情報量の低下を最小限にする．

と は，識別力の高い項目だけを選ぶと，「漢字の読みと仲間はずれの選択」に偏ってしまう．さらに項目の困難度がやさしいか，難しいかのどちらかになり，中

程度の困難度を持つ項目が選定からもれてしまう。

は、DAT 患者と VD 患者を識別する為のコンサイス版を目指した。

以上の選定基準によって、A1、A2、B1、B2、C、D と 6 種類のコンサイス版を作成した。A1 が第 1 因子負荷量の大きい順、B1 が項目情報関数の高い順である。そして、A2 と B2 は、時間を短縮するために 1 つの下位項目から 1 つの下位検査項目得点を選ぶという第 1 の選定基準に従い、A1 と B1 から「迷路の所要時間」を除いたものである。C は、項目困難度 (0.5 ~ 3.8) と項目識別力 (1.1 ~ 2.1) を 3×2 の 6 コの範疇に分け、各範疇から識別力の最も高い下位検査項目得点を 1 つずつ選定した。D は、判別分析における構造行列の相関係数の高い順に下位検査項目得点 1 コから 10 コまでの組み合わせを作り、その中で判別率の最も高かった組み合わせである。選定の結果、コンサイス版の下位検査項目の数は、A1 が 8、A2 が 7、B1 が 8、B2 が 7、C が 5、D が 7 となった。

結果

6 種類のコンサイス版の素点合計得点の平均値は、11.3 から 17.4 だった。病型別の素点合計得点の平均値は、DAT が平均 13.2 点、VD が 16.5 点であった。コンサイス版による素点合計得点の平均値における DAT と VD 間には、A1 が $p = 0.028$ 、A2 が $p = 0.026$ 、B1 が $p = 0.035$ 、B2 が $p = 0.034$ 、C が $p = 0.008$ 、D が $p = 0.003$ の有意確率で差があった。

作成したコンサイス版を検査所要時間、項目特性、テスト情報関数、信頼性、全検査とコンサイス版との推定尺度値の比較、全検査とコンサイス版での素点合計得点の比較、全検査による推定尺度値とコンサイス版の素点合計得点との比較、回帰係数、判別率の点から評価した。評価の結果を表 4 に示す。

- ・ 検査所要時間は、全検査施行の場合の平均 20 分から平均 5 分に短縮された。
- ・ 平均項目識別力は、1.32 から 1.86 で 項目識別力と 項目困難度の基準によったコンサイス C が 1.86 と全検査の 1.18 を大きく上回っていた。平均項目困難度は、構造行列相関係数の基準によったコンサイス D が 0.81 と全検査に最も近かった。
- ・ 最大テスト情報関数は 6.52 ~ 10.32 で、項目情報関数の基準によったコンサイス B1 が最も高かった。平均テスト情報関数は 3.73 から 5.94 で、B1 が最も高かった。平均テスト情報関数を項目数で割って得た平均項目情報関数は 0.59 から 0.75 と、全てのコンサイス版が全検査より高かった。

- ・ テスト信頼性係数は 0.79 から 0.86 と ,全てのコンサイス版で高かった .また ,Cronbach' も 0.67 から 0.79 と ,全てのコンサイス版で高かった .
- ・ 全検査による推定尺度値と ,コンサイス版による推定尺度値との相関係数は , $r = 0.86 \sim 0.91$ ($p < .001$) と ,全てで高かった .
- ・ 全検査の素点合計得点とコンサイス版の素点合計得点との相関係数は , $r = 0.92 \sim 0.96$ ($p < .001$) と ,全てのコンサイス版で高かった .
- ・ 全検査による推定尺度値とコンサイス版の素点合計得点との相関係数は , $0.92 \sim 0.96$ ($p < .001$) と十分高かった .この結果を踏まえ ,回帰式によりコンサイス版の素点合計得点から全検査による推定尺度値を推定した .全検査による推定尺度値のコンサイス版の素点合計得点への回帰係数は , 0.13 から 0.17 であった .
- ・ 6 種類のコンサイス版による判別率は 62.0% から 64.5% で ,コンサイス A2 が最も高かった .

これらの結果から ,痴呆の重症度診断には全てのコンサイス版が ,病型分類にはコンサイス A2 が適切と考えられる .

2) シングル版

項目の選定

検査所要時間のさらなる短縮を目指したシングル版は ,単一の下位検査課題のみを用いるものである .シングル版の作成には ,最大WMデマンド数 ,第 1 因子負荷量 ,項目識別力 ,項目情報関数 ,および構造行列の相関係数の点から吟味した .

最大WMデマンド数 7 以上の項目は ,事物の階層構造的分類 ,おはじきの数の再現 3 桁 ,漢字の仲間はずれの選択 ,迷路であった .

第 1 因子負荷量の 0.75 以上の項目は ,事物の階層構造的分類 ,おはじきの数の再現 ,漢字の仲間はずれの選択 ,迷路 2 であった .

項目識別力が 1.3 以上の項目は ,事物の階層構造的分類 1 ,おはじきの数の再現 3 桁 ,漢字の読みと仲間はずれの選択 ,迷路 1 , 3 であった .

項目情報関数が 0.5 以上の項目は ,事物の階層構造的分類 ,おはじきの数の再現 3 桁 ,漢字の仲間はずれの選択 ,迷路 2 , 3 であった .

構造行列の相関係数が 0.2 以上の項目は ,事物の階層構造的分類 1 ,おはじきの数の計数 3 桁と再現 3 桁 ,漢字の仲間はずれの選択 ,迷路であった .

以上の結果から、シングル版を「事物の階層構造的分類 1」「おはじきの数の再現 3 桁」「漢字の仲間はずれ 村」「迷路 2」の 4 コに決定した。

結果

シングル版の検査所要時間は約 3 分である。

シングル版の重症度診断の正確度の検証については、項目反応理論が適用できないので、1 項目からなる全検査の素点合計得点とシングル版の素点合計得点との相関係数によって確認した。項目反応理論の適用ができないのは、BILOG を用いて項目パラメータを推定するためには 3 コ以上の項目数を必要とするという制約があるからである (Mislevy, 1900)。

結果を表 5 に示した。全検査の素点合計得点とシングル版の素点合計得点との相関係数は、 $r = 0.77 \sim 0.79$ ($p < .001$) と全てで高かった。また、全検査による推定尺度値とシングル版の素点との相関係数は、 $0.64 \sim 0.74$ ($p < .001$) と高かった。そこで、シングル版の素点から全検査による推定尺度値のための回帰式を求めた。そのうち A (事物の階層構造的分類) の回帰係数は 0.97 と非常に高かった (表 5)。図 3 に、シングル版の素点から全検査による推定尺度値を推定するための回帰直線を示した。

判別率は 52.9% ~ 62.8% で、シングル版 D (迷路 2) が最も高かった。

これらの結果から、痴呆の重症度診断には A、病型分類には D が適切と考えられる。

3) 既存の評価法とコンサイス版、およびシングル版との比較

簡易検査法の痴呆の重症度診断テストとしての有用性を確認する為に、既存の評価法と比較し、所要時間、重症度の判定度、信頼性の点から検証する。

テスト情報関数が最大となる推定尺度値は、コンサイス A1 が 0.13、A2 が 0、B1 が 0.38、B2 が 0、コンサイス C が - 1.63、そしてコンサイス D が - 0.25 だった。この結果を、図 2 に示した行動評価得点と全検査による推定尺度値との関係に照らすと、痴呆症状が軽度から中等度の患者にはコンサイス A、A2、B、B2、やや高度の患者にはコンサイス D、そして高度の患者にはコンサイス C を施行することにより、より多くの情報量が得られることを示している。

なお、コンサイス版の素点合計得点、およびシングル版の素点と FAST、CDR、HDS-R、MMSE、および ADAS-Jcog との相関係数の値は、全てのコンサイス版で高く、高い併存的妥当性、シングル版ではやや低く、中程度の並存的妥当性が確認された。

あとがき

本研究は、老年痴呆、特に DAT の病態が WM 容量の著しい減少に起因するという仮説に基づき、知能研究の手法として心理学の中で確立されてきた心理測定学的方法と WM 理論という認知心理学的アプローチを組み合わせることによって、新しい痴呆診断のための神経心理学的検査（TKW 式検査）の作成を企図した。

予備的な二つの研究の結果により、理論的基盤に立脚した項目を中心に構成された TKW 式検査が、痴呆診断により有効な新しい検査としての条件を備えるようになった。その特徴の一つは、WM 理論への当てはまりが良く、TKW 式検査の測定対象が WM 容量であることが裏づけられたことである。

二つ目の特徴は、心理測定学の成果である項目反応理論を適用したことである。TKW 式検査が均質テストであるという項目反応理論適用の前提条件が確認され、測定対象である単一次元が WM 容量であることが明らかになった。そして、項目反応理論を適用したことにより、全検査項目の一部分、あるいは単一の検査項目の取り扱いが可能となり、6 種類のコンサイス版と 4 つのシングル版の作成に成功した。施行時間が約 20 分から約 3～5 分に短縮されたこれらの簡易検査法は、全検査による推定尺度値と高い相関を示した。さらに、日常生活動作能力との相関係数が高く、痴呆の重症度診断に有用であることが確認できた。そして、軽度から中等度の痴呆患者にはコンサイス A、A2、B、B2 が、やや高度の痴呆患者にはコンサイス D、そして高度の痴呆患者にはコンサイス C というように、痴呆の症状に合わせてコンサイス版を施行することにより、最大限の情報量を得ることが可能となった。下位検査項目 1 コからなるシングル版 A は、全検査による推定尺度値を推定するための回帰係数が 0.97 とほぼ 1 であることから、得られた得点から 1.2 を引き算するだけで全検査による推定尺度値を得ることができ、臨床現場における有用性は高いと思われる。項目反応理論の適用により、臨床現場における時間的な制約と、痴呆の重症診断のためには、時間に縛られない力量検査がより適切であるというジレンマが解消された。

一般に、心理検査による病型分類については、記憶障害が前面に出る DAT に比べ、VD は感情障害、人格障害、あるいは性格変化が前面に出るため、HDS-R や MMSE による鑑別診断は困難とされる（祖父江ほか、1999）。本研究においても、素点合計得点による判別率は、HDS-R が 55.0%、TKW 全検査が 58.7%であった。しかし、項目反応理論の特徴を生かし、検査得点を構成する下位検査項目の単位で扱うことにより、TKW 全検査の判別率を 87.6%まで高めることが可能となった。病型診断のスクリーニングに十分役立つという

のが本検査の三つ目の特徴である。下位検査項目単位での病型の分類については、「事物の階層構造的分類」と「漢字の仲間はずれの選択」が有効と考えられたが、今後さらに各病型の被検査者の下位検査項目に対する反応パターンを収集することにより検証していきたい。

重症度診断能力については、日常生活動作能力との相関係数が $r = -0.81$ から -0.85 と高く、痴呆の重症度診断テストとしての有効性が検証された。しかし、痴呆の経過については、短期間の追跡結果しか報告できなかった。今後さらに追跡を続けることにより痴呆の重症度診断テストとしての有効性の検討をしていきたい。

最後に、本論文の目次、研究業績一覧、および引用文献を付した。

表5 シングル版の全検査との比較

	シングル A	シングル B	シングル C	シングル D
下位検査項目名	分類 1	再現 3桁	仲間はずれ 村	迷路 2
全検査の素点合計得点との相関係数 ^{*1}	0.77	0.78	0.79	0.77
全検査による推定尺度値との相関係数 ^{*2}	0.74	0.64	0.74	0.72
推定尺度値との回帰式	$\hat{y} = 0.97x - 1.2$	$\hat{y} = 0.27x - 1.1$	$\hat{y} = 0.42x - 1.3$	$\hat{y} = 0.44x - 0.8$
判別率 (%)	56.2	52.9	66.1	65.3

*1：全検査（36コの下位検査項目得点）による素点合計得点とシングル版の素点との相関係数
有意水準は、全て $p < .001$ であった。

*2：全検査による推定尺度値とシングル版の素点との相関係数
有意水準は、全て $p < .001$ であった。

本論文の目次

第1章 序論

1-1 本研究の位置づけ

1-2 痴呆の病態の解釈

1-3 本研究の目的

1-4 本論文の構成

第2章 神経心理学と痴呆症状の神経心理学的検査

2-1 神経心理学と神経心理学的検査

2-2 痴呆とは

2-3 主な痴呆評価法，および本研究で用いた痴呆評価法の特徴

2-4 Lezakの遂行機能理論による痴呆症状の理解

第3章 痴呆症状と作動記憶理論，および知能因子構造論との関係

3-1 作動記憶理論

3-1-1 Logieによる作動記憶理論の7つの時代

3-1-2 Baddeleyの作動記憶モデル

3-1-3 情報処理低下についての作動記憶モデルによる説明

3-1-4 Caseの作動記憶モデル

3-2 作動記憶と流動性知能との関係

3-2-1 知能因子構造論

3-2-2 Horn-Cattellによる流動性・結晶性知能理論

第4章 テスト理論

4-1 古典的テスト理論

4-2 現代テスト理論

4-2-1 Guttmanの尺度化理論

4-2-2 項目反応理論

第5章 痴呆重症度診断のための神経心理学的検査（TKW式検査）の作成経緯

5-1 研究（予備研究1）

5-1-1 目的

5-1-2 対象

5-1-3 検査項目，施行手続き，および採点法

- 5-1-4 結果と考察
- 5-2 研究（予備研究2）
 - 5-2-1 目的
 - 5-2-2 対象
 - 5-2-3 結果と考察
- 5-3 作動記憶理論に基づく下位項目の決定
- 第6章 TKW式検査の施行とその結果，および概念的妥当性の検証
 - 6-1 対象と施行手続き
 - 6-2 結果
 - 6-2-1 所要時間と検査得点
 - 6-2-2 因子分析によるTKW式検査の均質性についての検証
 - 6-3 TKW式検査が作動記憶理論に立脚していることを確認するための分析
 - 6-3-1 最大作動記憶デマンド数による理論的項目困難度との適合度の検証
 - 6-3-2 Guttmanの尺度化理論との適合度の検証
 - 6-3-3 項目反応理論の適用による推定尺度値との適合度の検証
 - 6-3-4 流動性知能という観点からの妥当性の検証
 - 6-4 病型の分類
- 第7章 既存の評価法とTKW検査との比較
 - 7-1 対象と施行手続き
 - 7-2 結果
 - 7-3 TKW式検査と既存の評価法との比較による併存的妥当性の検証
 - 7-4 痴呆の重症度診断テストとしての予測的妥当性と信頼性の検証
- 第8章 TKW式検査の訓練効果
 - 8-1 痴呆患者に対する訓練効果についての先行研究
 - 8-2 目的
 - 8-3 方法
 - 8-3-1 対象
 - 8-3-2 訓練課題
 - 8-3-3 訓練手続き
 - 8-4 結果

第9章 臨床場面における実用化

9-1 目的

9-2 コンサイス版

9-2-1 項目の選定

9-2-2 対象

9-2-3 結果

9-3 シングル版

9-3-1 項目の選定

9-3-2 結果

9-4 既存の評価法とコンサイス版，およびシングル版との比較

第10章 あとがき

図表一覧

研究業績一覧

引用文献

謝辞

研究業績一覧

修士論文

作動記憶理論に基づく老年痴呆の神経心理学的診断検査の作成と妥当性の検証．
1997.3.15. 早稲田大学大学院教育学研究科．

論文

作動記憶理論による学習困難の分析：学業不振と老年痴呆の場合．2001．早稲田大学大学院教育学研究科紀要，別冊，8，85-96．

（共著）作動記憶理論と項目反応理論に基づく痴呆患者の神経心理学的検査（T-K-W 式検査）：検査法作成の経過．精神神経学雑誌，2002，104，690-709．

（共著）作動記憶理論と項目反応理論に基づく痴呆患者の神経心理学的検査（T-K-W 式検査）：検査法作成の経過．精神神経学雑誌，審査中．

学会発表

（共著）作動記憶理論に基づく老年痴呆の神経心理学的診断検査の作成と妥当性の検証．
日本教育心理学会第 39 回総会発表論文集，1997，p. 363．

神経心理学的診断テスト結果から見た老年痴呆患者の作動記憶容量と一般知能因子との関係．第 3 回 JCOM 発表論文集，1999，p. 11．

（共著）老年痴呆の神経心理学的診断テストの理論的裏付け．日本教育心理学会第 41 回総会発表論文集，1999，p.467．

（共著）老年痴呆の神経心理学的診断検査得点の分析 作動記憶理論と項目反応理論との関連性．日本教育心理学会第 42 回総会発表論文集 2000，p.715．

（共著）老年痴呆患者への訓練効果．日本心理学会第 64 回発表論文集，2000，p.254．

（共著）作動記憶理論と項目反応理論に基づく痴呆の為の神経心理学的検査（T-K-W 検査）の開発：痴呆重症度診断，および病型診断の為の用途別コンサイス版の構成．日本心理学会第 67 回発表論文集，2003，p. 181．

引用文献

- The American Psychiatric Association 編. 高橋三郎, 大野裕, 染矢俊幸訳 (1982) DSM-III 精神障害の分類と診断の手引. 東京, 医学書院.
- The American Psychiatric Association 編. 高橋三郎, 大野裕, 染矢俊幸訳 (1988) DSM-III-R 精神障害の分類と診断の手引. 第2版. 東京, 医学書院.
- The American Psychiatric Association 編. 高橋三郎, 大野裕, 染矢俊幸訳 (1995) DSM-IV 精神疾患の分類と診断の手引. 東京, 医学書院.
- The American Psychiatric Association 編. 高橋三郎, 大野裕, 染矢俊幸訳 (2002) DSM-IV-TR 精神疾患の分類と診断の手引. 東京, 医学書院.
- Backman, L., et al. (1992) Memory training and memory improvement in Alzheimer's disease: Rules and exceptions. *Acta Neurologica Scandinavica*, Supplement 139, 84-89.
- Baddeley, A. D. (1986) *Working memory. Oxford Psychology Series*, 11, 224-253, Oxford, Oxford Univ. Press.
- Baddeley, A. D. (1997) *Human Memory: Theory and Practice (Revised Edition)*. Hove, East Sussex, Psychology Press.
- Baddeley, A. D. (1999) *Essential of human memory*. 251-272, Hove, East Sussex, Psychology Press.
- Baddeley, A. D. & Hitch, G. J. (1974) Working Memory. In G. Bower (Ed.) *Recent advances in learning and motivation*, 8, 47-90, New York, Academic Press.
- Baddeley, A. D., Logie, R., Bressi, S., Della Salla, S. & Spinnler, H. (1986) Dementia and working Memory. *The Quarterly Journal of Experimental Psychology*, 38A, 603-618.
- Baddeley, A. D., Bressi, S., Della Sala, S., et al. (1991) The decline of working memory in Alzheimer's disease: A longitudinal study. *Brain*, 114, 2521-2542.
- Baltes, P. B., & Smith, J., Staudinger, U. M. (1992) Wisdom and successful aging. In T. B. Sanderregger (Ed.). *Psychology and aging*. 123-167. Lincoln, Neb. University of Nebraska Press.
- Baltes, P. B., & Willis, S. L. (1982) Plasticity and Enhancement of Intellectual Functioning in old Age: Penn State's Adult Development and Enrichment Project (ADEPT). In F. I. M. Craik and S. Treub (Eds.) *Aging and cognitive processes*. 353-389, New York, Plenum Press.

- Becker, J. T. (1988) Working Memory and secondary memory deficits in Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, 10, 739-753.
- Brandt, J, and Rich, J.B. (1995) Memory Disorders in the Dementias. In A.D. Baddeley, B.A. Wilson, and F.N. Watts (Eds.) *Handbook of memory disorders*. 243-270, Chichester, John Wiley & Sons.
- Campione, J. L., Brown, A. L., Bryant & N. R. (1985) Individual difference in learning memory. In R. Sternberg (Ed.) *Human abilities: An information-processing approach*, 103-126. New York, W. H. Freeman.
- Carlesimo, G. A., Fadda, L., Lorusso, S., Caltagirone, C. (1994) Verbal and Spatial memory spans in Alzheimer's and multi-infarct dementia. *Acta Neurologica Scandinavica*, 89, 132-138.
- Carroll, J.B.(1997)The three-stratum theory of cognitive abilities. In D. P. Flanagan, J.L. Genshaft, & P.L.Harrison (Eds.) *Contemporary intellectual assessment: Theories, tests, and issues*. 122-130. New York, Guilford.
- Case, R.(1980)The Underlying Mechanism of Intellectual Development. In J. R. Kirby, J.B. Biggs (Eds.) *Cognition, Development, and Instruction*. 5-37, South Wales, Academic Press.
- Chertkow, H. & Bud, D.N. (1990). Semantic Memory Loss in Dementia of AD Type. *Brain*, 113, 397-417.
- Craik, F.I.M., Morris, R.G. and Gick, M.L. (1990) Adult age differences in working memory. In G. Vallar and T. Shallice (Eds.) *Neuropsychological impairments of STM*. 247-267, New York, Cambridge University Press.
- D'Agostino, P. R., O'Neill, B. & Paivio, A (1977) Memory for Pictures and words as a function of level of processing: Depth or dual coding? *Memory and Cognition*, 5, 252-256.
- Denney, N. W., & Heidrich, S. M., (1990). Training effects in Raven's Progressive Matrices in young, middle-aged, and elderly adults. *Psychology and Aging*, 5, 144-145.
- Diesfeldt, H.F.A. (1985) Verbal fluency in senile dementia: An analysis of search and knowledge. *Archives of Gerontology and Geriatrics*, 4, 231-239.
- Embretson, S. E. (1994) Latent trait theory. In R. J. Sternberg (Ed.), *Encyclopedia of Human Intelligence*. 644-647, New York, Macmillan.
- Engle, R. W. et al.(1999a)Working Memory, Short-Term Memory, and General Fluid Intelligence: A Latent-Variable Approach. *J. of Experimental Psychology. General*, 128, 309-331.

- Engle, R. W. et al. (1999b) Individual differences in working memory capacity and what they talk us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.) *Models of working memory: Mechanisms of active maintenance and executive control*. 102-134, New York, Cambridge Univ. Press.
- Erdelyi, M. H. & Kleinbard, J. (1978) Has Ebbinghaus decayed with time? The growth of recall (hyperamnesia) over days. *J. of Experimental Psychology*. 4, 275-289.
- Fillenbaum, G. G., Heyman, A., Wilkinson, W. E. & Haynes, C. S. (1987) Comparison of two screening test in Alzheimer's Disease. The correlation and reliability of the Mini-Mental State Examination and the Modified Blessed Test. *Arch Neurol*, 44, 924-927.
- Finger, S. 河内十郎訳 (1998) 神経心理学の歴史 . In D. W. Zaidel , 河内十郎監訳. 神経心理学 : その歴史と臨床の現状 . 19-34 , 東京 , 産業図書株式会社.
- Flavell, J.H. (1978) Comments on Case's paper. In R.S. Siegler (Ed.) *Children's thinking: What develops?* Hove, East Sussex, LEA.
- Flynn, T. & Storndt, M. (1990) Supplemental group discussions in memory training for older adults. *Psychology and Aging*, 5, 178-181.
- Folstein, M. F., Folstein, S. E., McHugh, P. R. (1975) "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatric Research*, 12, 189-198.
- 福沢一吉, 辰巳格, 笹沼澄子 (1988) 痴呆患者における語想起の特徴について. *失語症研究*, 8, 243-250.
- Gatz, M. et al. (1998) Empirically Validated Psychological Treatments for Older Adults. *J. of Mental Health and Aging*, 4, 9-46.
- Green, B.F. (1956) Method of Scalogram analysis using summary statistics. *Psychometrika*, 21, 79-88.
- Guttman, L. (1947) The Cornell Technique for scale and intensity analysis. *Educational and Psychological Measurement*, 7, 247-279.
- Guttman, L. (1900) The basis for scalogram Analysis. In S. A. Stouffer et al. (Eds.) *Measurement and prediction: Studies in social psychology in World War* . 4, 60-90, N. J. Princeton, Princeton University Press.
- Hecaen, H. (1978) *Human neuropsychology*. Canada, John Wiley & Sons, Inc.
- K. M. Heilman, E. Valens (Eds.) 杉下守弘監訳 (1995) 臨床神経心理学. 東京, 朝倉書店.

- Herlitz, A. & Viitanen, M. (1991) Semantic organization and verbal episodic memory in patients with mild and moderate AD. *J. of Clinical and Experimental Neuropsychology*, 13, 559-574.
- Hill, R.D., Storandt, M. & Simeone, C. (1990) The effects of memory skills training and incentives on free recall in older learners. *J. of Gerontology*, 45, 227-232.
- Hofland, B.E., et al. (1981) Fluid intelligence performance in the elderly: Intraindividual variability and conditions of assessment. *J. of Educational Psychology*, 73, 4, 573-586.
- Horn, J.L. (1982) The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In F.I.M. Craik and S. Treub (Eds.) *Aging and cognitive processes*. 237-278, New York, Plenum Press.
- Horn, J. L. (1989) Models of Intelligence. In R. Linn (Ed.) *Intelligence: measurement, theory, and public policy*. 29-73, Urbana-Champaign, University of Illinois Press.
- Horn, J.L. & Hofer, S.M. (1992) Major Abilities and Development in the Adult Period. In R.J. Sternberg & C.A. Berg, (Ed.) *Intellectual development*. 44-99, New York, Cambridge University Press.
- Hughes, C.H., Berg, L., Danziger, W. L., Coben, L.A., and Martin, R. L. (1982) A New Clinical Scale for the Staging of Dementia. *Brit. J. Psychiat.* 140, 566-572.
- Hwang, J.P. et al. (2000) Repetitive phenomena in dementia. *International Journal of Psychiatry in Medicine*, 30,
- 本間昭 (1993a) 痴呆の客観的評価法, 脳神経, 45, 119-1133.
- 本間昭 (1993b) 痴呆の新しい評価法, 老年期痴呆, 7, 111-118.
- 本間昭 (1996) 在宅痴呆性老人の実態. 東京都福祉局 平成 7 年度高齢者の生活実態及び健康に関する専門調査結果報告書. 財団法人 静岡総合研究機構.
- 本間昭 (1992) アルツハイマ - 型痴呆における臨床治験のガイドライン. 老年精神医学雑誌, 3, 905-910.
- 本間昭, 福沢一吉, 塚田良雄ほか (1992b) Alzheimer's Disease Assessment Scale (ADAS) . 日本版の作成. 老年精神医学雑誌, 3, 647-655.
- 池田央 (1994) 現代テスト理論. 東京, 朝倉書店.
- 今栄国晴 (1968) 各種言語材料の共通測度としての学習容易性評定 . *Jap. J. of Psychol.* 39, 241-245.

- Kaiser, H.F. (1960) The Application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141-151.
- Kawata, T., Namiki, H., Yamamoto, J., et al. (1979) Experimental verification of Mental-Space theory and its problem at issue. *Studies in Sociology, Psychology, and Education*, 19, 91-97, Keio University.
- Klatzky, R.L.(1984) *Memory and Awareness: An Information-Processing Perspective*. New York , W.H. Freeman and Company.
- Kyllonen, P.C. (1996) Is Working Memory Capacity Spearman's g. In I. Dennis & P. Tapsfield (Eds.) *Human abilities: Their nature and measurement*. 49-75, New jersey, Lawrence Erlbaum Associates, Inc.
- 海保博之 (1975) 漢字意味情報抽出過程. 徳島大学学芸紀要, 24, 1-7.
- 鹿島晴雄 (1996) 神経心理学的検査法：神経心理学的検査における定量的アプローチと定性的アプローチ. 臨床精神医学, 12 月増刊号, 139-145.
- 鹿島晴雄他 (1999) 認知リハビリテーション. 155-175, 東京, 医学書院.
- 加藤元一郎 (1995) 注意障害 - 臨床的理解とリハビリテーション. In 江藤文夫, 原寛美, 板東充秋編 *Clinical Rehabilitation 別冊：高次機能障害のリハビリテーション*. 22-29, 東京, 医歯薬出版.
- 加藤元一郎 (2003) アルツハイマー病の神経心理学. 第 17 回老年期痴呆研究会口頭発表 .
- 加藤元一郎, 鹿島晴雄 (1996) 神経心理学的検査法：遂行機能. 171-179, 東京, 臨床精神医学.
- 加藤伸司 (1996) テスト式知的機能検査法とその問題点：B.欧米で開発された検査法 *Dementia*, 10, 259-277.
- 金子満雄 (1991a) 痴呆早期診断の為のスクリーニングテスト. 老年期痴呆, 4, 41-48.
- 金子満雄 (1991b) 早期老年痴呆の診断と治療. 老年期痴呆研究会誌, 5, 102-104.
- 金子満雄, 山本貴道, 田中敬生, 高槻絹子ほか (1989) 早期老年痴呆に対する脳活性化訓練の効果：神経心理テストによる評価, および追跡結果について. 日本医事新報, 3397, 28-31.
- 北尾倫彦他 (1977) 教育漢字 881 字の具体性, 象形性, 及び熟知性. *Jap. J. of Psychol.* 48, 105-111.
- 黒田洋一郎 (1998) アルツハイマ - 病 . 東京, 岩波書店.

- 小林敏子 (1996) 日常生活動作能力評価尺度とその問題点. *Dementia*, 10, 297-311.
- Lezak M.D. (1982) The problems of assessing executive functions. *International Journal of Psychology*, 17, 281-297.
- Lezak M.D. (1995) *Neuropsychological Assessment*, 3rd ed. 650-685. New York. Oxford University Press.
- Linn, R. L.池田央, 藤田恵璽, 柳井晴夫, 繁桝算男訳 (1992) 教育測定学上下. 原著第 3 版. 東京, 学習評価研究所.
- Linn R. T., Wolf, P. A., Bachman, D. L., et al. (1995) The 'preclinical phase' of probable Alzheimer's disease. A 13-year prospective study of the Framingham cohort. *Arch Neurol* 52, 485-490.
- Logie, R. H. (1996) The seven ages of working memory. In John Richardson, et al. (Eds) *Working Memory and Human Cognition*. 31-65, Oxford University Press, Inc.
- Luria, A.R.保崎秀夫監修, 鹿島晴雄訳 (1978) 神経心理学の基礎 - 脳のはたらき. 東京, 医学書院.
- Mackenzie, D. M. et al. (1996) Brief cognitive screening of the elderly: A comparison of the Mini-Mental State Examination, Abbreviated Mental Test and Mental Status Questionnaire. *Psychological medicine*. 26, 427-430.
- McGrew, K.S.(1998)*The intelligence test desk reference: Gf-Gc cross-battery assessment*. Boston, Allyn & Bacon.
- McKhann, G. et al. (1984) Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, 34, 939-944.
- Meiran N. et al. (1996) Diagnosis of Dementia: Methods for interrelation of scores of 5 neuropsychological tests. *Archives of Neurology*, 53, 1043-1054.
- Mislevy, R., Bock, R.D. (1900) BILOG 3. Item Analysis and Test Scoring with Binary Logistic Models. 2nd Edition. Mooresville, IN. Scientific Software, Inc.
- Mohs, R C. et al. (1997) Development of cognitive instruments for use in clinical trials of antidementia drugs: Additions to the Alzheimer's disease assessment scale that broaden its scope. *Alzheimer Disease and Associated Disorders*. 11, SUPPL. 2, S13-S21.

- Mohs, R.C., Rosen, W.G., Davis, K.L. (1983) The Alzheimer's Disease Assessment Scale: An instrument for assessing treatment efficacy. *Psychopharmacology Bull*, 19, 448-450.
- Morris, R.G. & Kopelman, M.D. (1986) The Memory Deficits in Alzheimer-Type Dementia: A Review. *The Quarterly Journal of Experimental Psychology*, 38A, 575-602.
- Morris, R. G. & Kopelman, M. (1994) The memory deficits in Alzheimer type dementia: A review, *Neuropsychology*, 8, 544-554.
- 三宅晶 (2000) ワーキングメモリー: 過去, 現在, 未来. 芋阪直行編著, 脳とワーキングメモリー. 311-324, 京都大学出版.
- 三宅晶, 斉藤智 (2001) 作動記憶研究の現状と展開. 心理学研究, 72, 336-350.
- 水野りか (1997) 漢字表記語の音韻処理自動化仮説の検証. 心理学研究, 68, 1-8.
- 森悦朗, 三谷洋子, 山鳥重 (1995) 神経疾患患者における日本版 MMSE テストの有用性. 神経心理学, 1, 82-90.
- Naglieri, J.A. (1997) Planning, Attention, Simultaneous, and Successive Theory and the Cognitive Assessment System: A New Theory-Based Measure of Intelligence. In D. P. Flanagan, J. L. Genshaft, P. L. Harrison (Eds), *Contemporary intellectual assessment: Theories, tests, and issues*. 247-267, New York, The Guilford Press.
- 中里克治 (1984) 老年期における知能と加齢. 心理学評論, 27, 247-259.
- 並木博 (1964) 幼児期における数量的思考に関する一実験. 慶應義塾大学大学院社会学研究科修士論文.
- 並木博 (1967) 語学ラボに対する態度尺度の構成. 慶應義塾大学語学視聴覚教育研究室紀要乙, 40-44.
- 並木博 (1982) 操作と作業記憶能力. In 波多野完治 (監), 天岩静子 (編) ピアジェ派心理学の発展 ピアジェ双書第5巻, 171-198, 東京, 国土社.
- 並木博 (1990) 認知機能の発達と学習. In 坂本昂 (編) 教職課程講座 2, 発達と学習. 79-101, 東京, ぎょうせい.
- 並木博 (1992) 発達的一般的原理, 大日向達子 (編) 発達心理学, 20-34, 朝倉書店.
- 並木博 (1993) 展望 教授・学習研究における ATI パラダイムと適性理論. 教育心理学年報, 32, 117-127.
- 並木博 (1996) 知能の生涯発達, 総合リハビリテーション, 24, 739-744.
- 並木博 (1997) 個性と教育環境の交互作用: 教育心理学の課題. 東京, 培風館.

- 並木博 (1999) 加齢に伴う知能の変化．東清和(編), エイジングの心理学, 31-64, 東京, 早稲田大学出版部.
- 並木博, 内藤俊史, 安岡龍太 (1976) ピアジェと測定：数量的思考と道徳的判断の場合，慶應義塾大学大学院社会学研究科紀要, 16, 1-19.
- 並木博, 篠原幸人, 山本正博, 米倉康江 (2002) 作動記憶理論と項目反応理論に基づく痴呆患者の神経心理学的検査(T-K-W 式検査): 検査法作成の経過 .精神神経学雑誌 ,104 , 690-709 .
- 西本武彦 (1982) 記憶実験用 picture 刺激の標準化 . 早稲田心理学年報 , 14 , 55-75.
- 大津由紀雄編(1995)読みの過程の普遍性と言語特異性. 認知心理学 3 言語. 194-197, 東京, 東京大学出版.
- 大塚俊男 (1996) テスト式知的機能検査法とその問題点 . A.我が国で開発された検査法 *Dementia*, 10, 245-258.
- 大友賢二 (1996) 項目応答理論入門 . 東京, 大修館書店.
- Pascual-Leone, J. (1970) A mathematical Model for the transition rule in Piaget's developmental stages. *Acta Psychologica*, 32, 301-345.
- Paivio, A., Smythe, P. C. & Yuille, J. C. (1968) Imagery versus meaningfulness of nouns in paired-associate learning. *Canadian J. of Psychology*, 22, 427-441.
- Peterson, R.C. et al. (1993) Longitudinal outcomes of patients with a mild cognitive impairment. *Ann Neurol* 43, 294.
- Robert, R.J., Hager, L.D. & Hertton, C. (1994) Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *J. of Experimental psychology: General*, 123, 374-393.
- Rogosa, D. (1982) A growth curve approach to the measurement of change. *Psychological Bulletin*, 92, 726-748.
- Salthouse, T. A. (1992) Working Memory as a Potential Mediator. In T.A. Salthouse (Ed.) *Mechanisms of Age-Cognition Relations in Adulthood*. 39-80. Hillsdale, New Jersey, Lawrence Erlbaum Associates Inc.
- Sasamura, S., Sakuma, N., and Kitano, K. (1992) Reading kanji without semantics: Evidences from a longitudinal study of patients with dementia. *Cognitive Neuropsychology*, 9, 465-486.
- Schaie, K.W. (1996) *Intellectual development in adulthood: The Seattle longitudinal study*. New York. Cambridge Univ. Press.

- Schaie, K.W. & Willis, S. L. (1986). Can Decline in Adult Intellectual Functioning be Reversed? *Developmental Psychology*, 22, 223-232.
- Small, B.J. & Backman, L. (1997) Cognitive Correlates of Mortality: Evidence From a Population-Based Sample of Very Old Adults. *Psychology and Aging*, 12, 309-313.
- Snodgrass, J.G. & Vanderwart, M. (1980) A Standardized set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity, and Visual Complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 2, 174-215.
- Solomon, P.R. et al. (1998) A 7 minute neurocognitive screening battery highly sensitive to Alzheimer's disease, *Archives of Neurology*, 55, 349-355.
- Spada, H. & Kluwe, R.H. (1980) Two Models of Intellectual Development and Their Reference to the Theory of Piaget. In R.H. Kluwe & H. Spada(Eds.)*Developmental models of thinking*, 1-31, New York, Academic Press.
- Sternberg, R. J. (1982) Education and intelligence. In Robert J. Sternberg (Ed.), *Handbook of human intelligence*. 493-585. Cambridge, Cambridge University Press.
- 坂爪一幸，本田哲三（1999）記憶障害の治療 認知リハビリテーション実践．臨床精神医学講座, 440-456, 中山書店.
- 坂野登（1984）教育心理学における神経心理学的接近．教育心理学研究, 32, 78-85.
- 佐藤恵子（1994）老人性痴呆の神経心理学的診断に関する考察．慶応義塾大学大学院社会学研究科修士論文．
- 祖父江逸郎（1999）脳血管性痴呆とアルツハイマ - 型痴呆の鑑別上の問題点．老年期痴呆研究会誌, 11, 89-119.
- 高島俊男（2001）漢字と日本人について．東京，文春新書.
- 豊田弘司（1998）精緻化の発達．心理学評論, 41, 1-14.
- 上野武治他，村田和香，吉田直樹（1996）痴呆老人のリハビリテーションについて．老年痴呆研究会誌, 9, 44-48.
- Warrington, E. K. (1975) The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, 27, 635-657.
- World Health Organization. 融道男，中根允文，小見山実監訳（1993）*ICD-10 精神および行動の傷害，臨床記述と診断ガイドライン*．東京，医学書院.

- Young, C.H., et al. (2000) The clinical applications of Mini-Mental State Examination in geropsychiatric inpatients. *The international Journal of Psychiatry in medicine*, 30, 277-286.
- 山鳥重 (1985) 神経心理学入門. 東京, 医学書院.
- 吉野文浩, 鹿島晴雄 (1996) 観察指揮行動評価尺度とその問題点. *Dementia*, 10, 179-196.
- 米倉康江 (1997) 作動記憶理論に基づく老年痴呆の神経心理学的診断検査の作成と妥当性の検証, 早稲田大学大学院教育学研究科修士論文.
- 米倉康江 (1999a) 神経心理学的診断テスト結果から見た老年痴呆患者の作動記憶容量と一般知能因子との関係. 第 3 回 JCOM 発表論文集, p. 11.
- 米倉康江, 深井英里子, 仲村禎夫, 並木博 (1997) 作動記憶理論に基づく老年痴呆の神経心理学的診断検査の作成と妥当性の検証, 日本教育心理学会第 39 回総会発表論文集, p.363.
- 米倉康江, 鎗田智子, 仲村禎夫, 並木博 (1999b) 老年痴呆の神経心理学的診断テストの理論的裏付け, 日本教育心理学会第 41 回総会発表論文集, p.467.
- 米倉康江, 野澤雄樹, 仲村禎夫, 並木博 (2000a) 老年痴呆の神経心理学的診断検査得点の分析: 作動記憶理論と項目反応理論との関連性. 日本教育心理学会第 42 回総会発表論文集, p.715.
- 米倉康江, 仲村禎夫, 並木博 (2000b) 老年痴呆患者への訓練効果. 日本心理学会第 64 回発表論文集, p.254.
- 米倉 康江 (2001) 作動記憶理論による学習困難の分析: 学業不振と老年痴呆の場合. 早稲田大学大学院教育学研究科紀要, 別冊, 8, 85-96.
- 米倉康江, 篠原幸人, 並木博 (2003) 作動記憶理論と項目反応理論に基づく痴呆の為の神経心理学的検査 (T-K-W 検査) の開発: 痴呆重症度診断, および病型診断の為の用途別コンサイス版の構成. 日本心理学会第 67 回発表論文集, p.181.
- Zaidel, D.W. 河内十郎監訳 (1998) 神経心理学 - その歴史と臨床の現状. 東京, 産業図書株式会社.
- Zanetti, O., Frisoni, G., De-Leo, D., Dello-Buono, A. et al. (1995) Reality orientation therapy in Alzheimer disease: useful or not? A controlled study. *Alzheimer Disease and Associated Disorders*, 9, 132-138.