

1. 論文種類・番号 研究資料 12016

2. 論文タイトル

和文表題 日本語版オペレーションスパンテストによるワーキングメモリの測定

英文表題 Assessment of working memory capacity with a Japanese version of the Operation Span Test.

3. 著者名・所属

小林 晃洋（専修大学大学院） / 大久保 街亜（専修大学）

Kobayashi, Akihiro (Senshu University Graduate school) / Okubo, Matia (Senshu University)

## 9. アブストラクト

The Operation Span Test (OSPAN) is widely used to assess working memory capacity. However, this instrument was rarely used to test Japanese participants because its task was not sufficiently difficult. The mean score for the original computerized OSPAN often reached a ceiling when Japanese participants were tested. In this study, we developed a computerized version of OSPAN for Japanese participants by increasing the task difficulty of the arithmetic procedures. The OSPAN scores were normally distributed and the mean score was approximately 50%. There were positive correlations between OSPAN scores and other scores of working memory measurements, such as a reading span test and a digit span test. These results suggest that the Japanese OSPAN is a reliable and valid measurement of working memory to test Japanese participants.

Keywords: working memory, Operation Span Test, Reading Span Test

## 10. 本文

人間の心がもつ最も優れた特長のひとつに、いつでもアクセスできるよう情報を心的に保持し、同時に情報の処理を行うという能力があげられる (Conway, Jarrold, Kane, Miyake, & Towse, 2008)。この能力はワーキングメモリ (Working memory) と呼ばれ、文章理解、計算などの問題解決、予定の設計など、高度な心的活動を行ううえで中心的な役割を担うとされる (e.g., Baddeley, 1986; Baddeley, 2007)。ワーキングメモリの容量は有限で、個人差や年齢による変動があるとされており、容量によって心的活動のパフォーマンスは変わると考えられている (e.g., Cowan, 2005; Conway et al., 2008)。

ワーキングメモリ容量の違いによるパフォーマンスの変化を検討するには、まず容量を測定することが不可欠である。そのため、これまでオペレーション・スパン・テスト (*Operation span test*, OSPAN, Turner & Engle, 1989) やリーディング・スパン・テスト (*Reading span test*, RSPAN, Daneman & Carpenter, 1980), カウンティング・スパン・テスト (*Counting span test*, Case, Kurland, & Goldberg, 1982) など様々な尺度が開発され用いられてきた。

これまで開発されてきたワーキングメモリ測定尺度の中で、代表的なものとして OSPAN と RSPAN があげられる。OSPAN は計算と単語の記銘を繰り返し、単語の再生成績によって容量を測定する尺度である。これに対し RSPAN は文章の音読と単語の記銘を繰り返し、単語の再生成績によって容量を測定する尺度である。どちらも計算や音読などの処理をしながら保持を行う能力を測定する尺度である。ただしワーキングメモリが処理・保持する対象は言語や視空間情報などさまざまであり (e.g., Baddeley & Hitch, 1974; Baddeley, 2000), 単一の尺度でワーキングメモリが持つ要素全てを測定する

ことは難しい。そのためワーキングメモリ容量を測定するために近年では **OSPAN** と **RSPAN** 等複数の尺度を用いることで測定精度を高める試みがなされている (e.g., Beilock & Carr, 2005; Tsai, Kang, & Peterson, 2010)。

わが国において **RSPAN** は苧阪・苧阪 (1994) により日本語版が開発され広く用いられている。しかしながら **OSPAN** については未だ日本語版の開発は試みられていない。この理由として、**OSPAN** が計算処理を行わせる課題を含んでおり、**RSPAN** と違い文化非依存的な尺度と考えられる点あげられる。しかしながら、海外で開発された **OSPAN** をそのままの手続きで日本人参加者に実施すると、参加者の得点が満点付近に集中し個人差の測定が困難になることが報告されてきた。たとえば小林・大久保 (2012) は日本人大学生に対し Unsworth, Heitz, Schrock, & Engle (2005) と同じ手続きで **OSPAN** を実施し、実験 1 では平均得点が 96% ( $SD = 1.9$ ) で得点の範囲が 85 - 100%、実験 2 では平均得点が 95% ( $SD = 2.1$ ) で得点の範囲が 77 - 100% というきわめて高い成績に偏る結果を示した。また大塚・宮谷 (2012, ポスターより) で実施された **OSPAN** (Unsworth, Redick, Heitz, Broadway, & Engle, 2009) では、平均得点が 89% ( $SD = 12$ )、計算の繰り返し回数を増やし処理の負荷を高めた **OSPAN** でも平均得点 81% ( $SD = 18$ ) という非常に高い成績を示し、平均得点  $+1SD$  で満点に達することがわかった。これらに対し Unsworth et al. (2005) では平均得点が 73.67% ( $SD = 18.27$ ) であり、日本人参加者の得点がきわめて高く偏っていることが示唆される。

オリジナルの **OSPAN** を日本人参加者に用いることが困難である可能性として、主に 2 点考えられる。ひとつは東アジア諸国と欧米諸国での数学能力の違い、もうひとつはアメリカと日本の研究におけるサンプルの違いである。数学能力の違いについて、これまで、**OSPAN**

が開発されたアメリカの学生に比べ、日本の学生は四則演算・代数・幾何・微積分などの様々な数学試験で高い成績を示すことが報告されている (Stevenson, Chen, & Lee, 1993; Stevenson, Lee, Chen, Stigler, Hsu, & Kitamura, 1990; Stevenson, Lee, & Stigler, 1986)。近年の国際比較 (*Programme for International Student Assessment*; PISA, OECD, 2010, p.15) でも、日本や台湾が最上位に位置するのに対しアメリカは中位である。この差は OSPAN で用いる四則演算で顕著である。たとえば Stevenson et al. (1993) は、加減算を習得する 6 - 7 歳児で 17.5%, 乗除算を習得する 9 - 10 歳児で 20% も日本人のほうが高い成績を示したと報告した。この傾向は成人においても同様であると予測される。これらの数学能力の違いによって、アメリカで開発された OSPAN の計算が日本人参加者には極めて簡単なため、処理への負荷として機能していない可能性が示唆される。大塚・宮谷 (2012) では計算と記銘の繰り返し回数を増やすことで、計算にかかる処理の負荷は変えないままでも高得点への集中が若干軽減されることを報告している。しかしながらこの改善でも平均得点 +1SD で満点に達する問題があり、個人差を幅広く測定するには不十分である。また処理への負荷が機能しないままでは認知能力の予測力も低下する (Engle, Tuholski, Laughlin, & Conway, 1999)。

もうひとつの可能性として、アメリカと日本の研究におけるサンプルの違いがあげられる。Unsworth et al. (2005) が行ったアメリカの研究では、18 - 35 歳の幅広い年齢にわたる一般市民に対し実験が行われた。これに対し小林・大久保 (2012) や大塚・宮谷 (2012) などの日本の研究では、実験参加者の多くが大学生または短大生であった。このため、日本の研究では、参加者の教育年数がアメリカの研究における参加者よりも比較的長い

ことで、オリジナルの OSPAN の成績が高くなった可能性が考えられる。ただし、オリジナルの OSPAN をアメリカの大学生に実施した多くの研究では、高い得点に偏らず、比較的広い範囲で個人差が測定されている (e.g., Jaeggi, Studer-Luethi, Buschkuhl, Su, Jonides, & Perrig (2010) では  $M = 74.11\%$  ( $SD = 18.56$ ), Shelton, Elliott, Matthews, Hill, & Gouvier (2010) では  $M = 58.87\%$  ( $SD = 20.72$ )). ここから日本における得点の偏りは、教育年数の違いのみでは説明できないことが示唆される。そこで本研究では、教育年数でなく数学能力の違いに着目し検討することとした。

OSPAN の特長は RSPAN や DSPAN など既存のワーキングメモリ尺度との相関が高いことに加え (Engle et al., 1999; Unsworth et al., 2005; Waters & Caplan, 1996), レーヴン漸進マトリクス (Raven's progressive matrices) や Scholarship aptitude test (SAT) などとの相関が高く、高度な認知能力に関して高い予測力をもつことである (e.g., Engle et al., 1999)。また OSPAN は RSPAN よりも繰り返しテストを行うことによる学習効果が低く (Unsworth et al., 2005; Waters & Caplan, 1996), 再検査による成績の変動が小さい (Unsworth et al., 2005; Waters & Caplan, 1996), すなわち再現性やテストーリテストの信頼性が高い尺度である。さらに数学不安が成績に影響を及ぼさない (Durette, 2011) ことなど、OSPAN はさまざまな利点をもつ。また前述のように RSPAN など複数の尺度を用いることで精度の高い測定が可能になる。したがって日本人にあわせた OSPAN の開発は、わが国におけるワーキングメモリ研究の発展に大きく貢献するものであると考えられる。

そこで本研究では、比較的高い数学能力を有していても、ワーキングメモリ容量の個人差を測定できるように、計算問題を改良した日本語版 OSPAN の開発を目的

とする。Unsworth et al. (2005) などが用いた計算問題について、本研究ではその難度を高める。OSPAN が開発されたアメリカでは数学教育に電卓を用いるのが一般だとされる (National Council of Teachers of Mathematics, 2000)。ここから OSPAN で用いるような四則演算を暗算で行う機会は日本人に比べ少ないと考えられる。これまでの OSPAN で用いられていた問題では、暗算や九九を行いやすいように、乗算すべき数字が視覚的に呈示されていた。たとえば Unsworth et al. (2005) は“(1 × 2) + 1 =?” (p.500) といった問題を用いた。この問題では乗算を行う数字が“(1 × 2)”と呈示されているため、九九を行いやすいと考えられる。これに対し本研究では、乗算すべき数字を視覚的には呈示せず、加算や減算を行った数字を心的に保持させ、その数字に対し乗算を行わせる問題とした。たとえば本研究では“(2 + 3) × 3 =?” といった問題を用いた。この問題は乗算と加算の組み合わせという点では Unsworth et al. (2005) と同じである。しかしながら、この問題を解くためには“(2+3)”を先に計算し、その答えである“(5)”を保持したうえで“(5 × 3)”を計算しなければならない。そのため、乗算すべき数字が視覚的に呈示されている Unsworth et al. (2005) に比べ暗算や九九への負荷は高まることが期待される。また本研究では Unsworth et al. (2005) において用いられていた 1 桁同士の除算だけでなく、九九の範囲内で 2 桁の除算も行う問題を用いた。たとえば本研究では除算を用いる問題として“(72 ÷ 8) - 3 =?” といった問題を用いた。2 桁の除算は 1 桁同士の除算に比べ負荷が高いうえ、呈示する除算の種類も多い。これを乗算の問題と組み合わせ様々な暗算を行わせることで、処理の自動化を抑制できると考えられる。さらにこれらの問題は、解答に必要な計算の回数が Unsworth et al. (2005) と変わらない。このため、計算に時間がかかりす

ぎるという問題も防げることが期待される。

そしてこの改良による成績を他の日本語版ワーキングメモリ尺度と比較することで妥当性を検討する。今回は、苧阪・苧阪 (1994) が開発し記銘材料が公開されている日本語版 RSPAN (苧阪, 2002) と、順唱 (*Forward digit span test*, DSPAN (F))・逆唱 (*Backward digit span test*, DSPAN (R)) からなる二つの数唱 (*Digit span test*, DSPAN) をワーキングメモリ尺度の比較対象とした。DSPAN のうち DSPAN (F) は、一般に短期記憶容量を測定する課題であるとされる (e.g., Engle, 2002)。一方 DSPAN (R) は、保持した数字を処理する課題のため、DSPAN (F) よりもワーキングメモリの処理に関わる課題であるとされる (Morris & Jones, 1990)。OSPAN, RSPAN は、DSPAN と同様に、どちらも記銘したターゲットの再生個数によって得点を算出する尺度である。もし本研究で実施した OSPAN が短期記憶の個人差を反映したものであれば、既存の尺度との相関は主に DSPAN (F) との正の相関として表れると考えられる。これに対しワーキングメモリの特徴である処理の要素を測定できていれば、その変動は RSPAN や DSPAN (R) との相関としても表れるだろう。

## 方法

### 参加者

参加者は専修大学に通う正常または矯正された視力をもつ学生 77 名 (女性 38 名, 男性 39 名) で、平均年齢 20.23 歳 ( $SD = 1.30$ ) であった。実験参加への同意を得た後、これまで OSPAN, RSPAN を受けた経験がないことを確認してからコンピュータの前に座り課題が行われた。なお DSPAN については WISC や WAIS (Psychological Corporation, 2002) などの知能検査でも採用されている課題のため、DSPAN の実施経験の有無によって参加者の選抜は行われなかった。OSPAN,



RSPAN, DSPAN (F), DSPAN (R) の順で課題が行われた。実験時間は 50 分程度であった。

### オペレーションスパンテスト (OSPAN)

Unsworth et al. (2005) の開発した OSPAN の計算問題と繰り返し回数 (スパン) を変更し実施された。Unsworth et al. (2005) の課題は Turner & Engle (1989) が開発した OSPAN をコンピュータ上で行えるようにしたものである (Figure 1)。本研究の OSPAN は、計算課題・記銘課題・再生課題・フィードバックからなった。各課題の詳細については後述する。OSPAN では画面に呈示される計算式の暗算・回答と、その後に呈示される 1 文字のアルファベットを記銘することが求められた。そして設定したスパンの回数だけ計算・記銘を繰り返したのち、呈示されたアルファベットを合図とともにすべて再生することが求められ、その後フィードバックが呈示されるものであった。実験では解像度が 1024 x 768 ピクセルの 17 インチ CRT モニタに課題を呈示し、マウスで反応取得が行われた。文字の大きさは全ての課題で 28 ポイントで呈示された。

Figure 1

計算課題・記銘課題 計算課題では画面中央に計算式が呈示された。計算課題については正答率 85% 以上を維持するよう、課題開始前に参加者に教示した。参加者は画面に呈示される計算式をできるだけ速く正確に暗算し、答えがわかったらマウスをクリックすることが求められた。クリックで反応すると画面が切り替わり、画面中央上部に数字が呈示され、画面中央下部に  が書かれたボタンと  が書かれたボタンが呈示された (Figure 1)。参加者はここで呈示された数字が、先ほど呈示された計算式の答えと同じ数字であったら ☐、そうでなければ ☐ をできるだけ速く判断しクリックすることが求められた (Figure 1 の例は ☐ が正解となる)。☐ か ☐ がクリックされると、選択に関するフィードバックはせずに

記銘課題にうつった。この課題では画面中央にアルファベット 1 文字が 800 ms 呈示され、参加者はこれを記銘することが求められた。記銘課題が終了すると直ちに次の計算課題が開始された。計算課題・記銘課題はスパンの回数だけ交互に繰り返された。スパンは 3 - 8 の 6 条件あり、参加者には現在行っている課題のスパン数は知らされなかった。また記銘課題で呈示されるアルファベットは 1 試行内で重複しなかった。

なお計算課題には Unsworth et al. (2005) と同様に制限時間が設けられており、制限時間内にマウスによる回答が得られなかった場合は記銘課題にうつった。制限時間は本試行前の練習試行での回答時間をもとに算出された (詳細は後述する)。

計算課題で呈示する計算式の出題形式は 2 種類あり、一つは九九の範囲内で乗算・除算を行う前に加算または減算を行わせる問題であった (e.g.,  $(2 + 3) \times 4 = ?$ )。もう一つは九九の範囲内で除算を行わせてから加算または減算を行わせる問題であった (e.g.,  $(18 \div 2) - 3 = ?$ )。計算課題において、誤答の場合は正解からランダムに 1 - 9、平均して 5 離れた値が呈示された。また、選択肢として呈示された値について、全試行数の半数が正答、残りの半数は誤答であった。記銘課題で呈示されるアルファベットは Unsworth et al. (2005) にならない “F, H, J, K, L, N, P, Q, R, S, T, Y” の 12 種類であった。

再生課題・フィードバック 計算課題・記銘課題をスパンの回数行ったのち、再生課題が行われた (Figure 1)。再生課題では記銘したアルファベットを順番に再生するよう教示する文章が画面最上部に呈示され、画面中央では各アルファベットが均等に配置され、各アルファベットの隣には選択用のボタンが呈示された。また画面下部にはリセットボタン (Reset) と決定ボタン (OK) が配置された。参加者は各アルファベットの隣にあるボタンを

順番にクリックすることで、記銘課題で呈示されたアルファベットを回答することが求められた。その際、回答した文字がわかるように、クリックした順番にボタン内に数字が表れ、また画面下部に選択した順番にアルファベットが文字列として表れた。回答し終えたら決定ボタンを押し再生課題は終了した。なお忘れた文字に関しては？ボタンで回答すること、また間違って回答してしまった場合はリセットボタンを押して再度選択できるよう教示文が呈示された。

再生課題後、フィードバック画面が 2000 ms 呈示された。この画面ではスパン数、再生課題で正しく回答された文字数、課題全体における現時点での計算課題正答率が呈示された。計算課題正答率が 85% を下回っていた場合、現時点での正答率とは別に画面下部に警告文が赤字で呈示された (Figure 1)。

OSPAN は 1 試行が計算課題、記銘課題、再生課題、フィードバックで構成された。このうち計算課題と記銘課題はスパンの数だけ 1 試行内で繰り返された。計算課題・記銘課題のスパンは 3 - 8 の 6 条件あり、ランダムな順で 6 試行行われた。この 6 試行を 1 ブロックとし、3 ブロック計 18 試行が行われた。そのため参加者は計算課題・記銘課題が全部で 99 問行われた。

練習試行・制限時間 練習試行は三つの段階からなっており、(a) 記銘と再生のみ行う、(b) 計算のみを行う、(c) スパンが 2 の計算課題・記銘課題、となっていた。まず計算は行わず記銘のみ 2 回繰り返し再生させる試行 (スパンが 2) を 3 試行行われた。次に計算の正誤判断のみを行う試行を 3 試行練習したのち 15 試行行われた。最後に計算課題・記銘課題でのスパンが 2 である試行が 3 試行行われた。このうち (b) の計算において、回答にかかった時間をもとに計算課題における制限時間が設定された。制限時間は Unsworth et al. (2005) になら

い、計算課題の平均回答時間  $+ 2.5SD$  であった。

### リーディングスパンテスト (RSPAN)

荳阪 (2002) の RSPAN を実施した。RSPAN は刺激文の音読と言葉の記銘を繰り返し行い、再生する合図が出た時点で記銘した言葉を順番に再生する課題であった。刺激文の呈示はコンピュータを用い、回答用紙に言葉を記述させた。

試行が開始されると、まずコンピュータ画面中央に刺激文が提示された。刺激文の一部にはターゲット語として下線が引かれていた。参加者は刺激文の音読を行いながらターゲット語を記銘することが求められた。音読が終了すると、実験者の操作によりただちに次の刺激文が提示され、参加者は同様に刺激文の音読とターゲット語の記銘が求められた。これを何度か繰り返すと、下線が引かれた言葉を順番に再生するよう教示する画面が提示された。参加者は、呈示された順にターゲット語を回答用紙に記入した。回答用紙には上部に再生の教示と同じ文章が記載されていた。用紙は 1 枚につき 5 試行分が割り当てられており、狭く書きにくいことのないようになっていた。また 1 試行ごとに線で区切られており、1 行でまとめて記入するようになっていた。記入し終わったら試行終了とし、次の試行は参加者の合図で開始された。

1 試行におけるスパン数は 2 - 5 の 4 条件あり、各条件につき 5 試行ずつ計 20 試行行われた。また本試行前にスパン 2 の練習試行が 2 試行行われた。なお RSPAN では OSPAN と異なり、スパン 2 を 5 試行ずつ行ったのちスパン 3 を 5 試行のように、スパンの小さい順に 5 試行ずつ行われた。刺激文は荳阪 (2002) と同じものが用いられた。各刺激文が明瞭に視認でき且つ 1 行に収まるように、文字の大きさは 17 インチモニタに 24 - 28 ポイントで呈示された。またターゲット語の下線は 2 ポイン

トの太線であった。なお課題開始前に、ある程度の声量で明瞭に読み上げること、刺激文が提示されたら直ちに音読を開始するよう教示された。課題中はこれらの教示に加え、不自然な読み方がなされた場合は随時注意された。またターゲット語は提示された順番に回答するよう教示され、最後に提示されたターゲット語から回答するのは避けるよう伝えられた。

### 数唱 (DSPAN)

DSPAN はまず DSPAN (F) が行われ、DSPAN (F) の課題が全て終了してから DSPAN (R) が行われた。刺激となる数字は、事前に録音したものがヘッドホンで聴覚呈示され、回答用紙に数字を記入させた。

DSPAN (F) では試行ごとに数字列が一度だけ聴覚呈示され、呈示後、聞こえた数字列が何であったかを回答用紙に記入するよう教示する画面が呈示された。音声刺激の呈示スピードは 1 秒につき数字 1 個であった。呈示された数字列は 3 - 9 桁の 7 種類あり、3 桁から 9 桁まで数字の小さい順に 2 試行ずつ計 14 試行行われた。また本試行前に数字列 3 桁の練習試行が 2 試行行われ、教示文は 28 ポイントの大きさに呈示された。

DSPAN (R) では聞こえた数字列を提示されたのとは逆順に回答用紙に記入することが求められた (e.g., 9, 1, 2 の数字列が提示された場合は 2, 1, 9 が正しい答えとなる)。先とは異なる課題であることを参加者に確認させるため、例として、課題前に実験者が 3 桁の数字を言い、逆順にするとどうなるか参加者に答えさせた。その後課題が開始された。数字列を聴覚呈示後、呈示された数字を逆の順番で書くよう太字で強調した教示が呈示され、回答用紙に記入させた。それ以外は DSPAN (F) と同じ条件で練習試行 2 試行、本試行 14 試行行われた。回答用紙は RSPAN と同様に上部に再生の教示と同じ文章が記載されていた。1 枚につき 5 試行分が割り当て

られており、狭く書きにくいことのないようにし、また 1 行で数字をまとめて記入するようになっていた。

### 結果

OSPAN において、計算問題の平均正答率は 93% であり、95% 信頼区間 (CI) は [92.21,93.79] であった。また計算問題が提示されてから正誤判断終了時までの平均反応時間は 4208 ms であり、95%CI は [3977, 4439] であった。

### 得点化

OSPAN は Unsworth et al. (2005) にならい、部分加点法 (*partial scoring method*) と完全加点法 (*absolute scoring method*) で得点を算出した。部分加点法は、呈示されたアルファベットを正しい位置で再生できた場合、再生できた分得点が加算される方法である。たとえば “P, J, K, T” が呈示され、“Q, J, H, T” と再生した場合、J, T が正しい位置で再生されたため 2 点加算される。これに対し完全加点法は、試行内で呈示されたアルファベットを呈示された順序ですべて再生できた場合に加点される方法である。たとえば OSPAN のスパン 7 の試行で 6 文字再生できても 0 点であり、7 文字を位置も正しく再生できて初めて 7 点が加算される。OSPAN は 99 点満点 (スパン 3 - 8 を 3 試行ずつ) であった。

RSPAN では Friedman & Miyake (2005) などを用いられている部分加点法と、先に述べた完全加点法で得点を算出した。この部分加点法では、呈示されたターゲット語を順不同で再生できた分だけ加点される方法である。たとえばスパン 4 の試行で 3 語再生されれば 3 点加算される。本研究では RSPAN は 70 点満点 (スパン 2 - 5 をそれぞれ 5 試行ずつ) であった。

DSPAN では、WAIS と同様に、提示された数字列を正しく再生できた場合のみ正解として文字数の分だけ加点した。これは一つの試行で同じ数字は一度しか用いな

いため、各試行ごとにすべての数字を回答すると部分加点法では正解となってしまうためである (e.g., 9, 4, 5, 1, 3, 7, 8, 2 と提示された試行で 1, 2, 3, 4, 5, 6, 7, 8, 9 と回答すると、部分加点法では 8 点が加算される)。DSPAN は DSPAN (F), DSPAN (R) とともに 84 点満点であった (スパン 3 - 9 をそれぞれ 2 試行ずつ)。得点分布・相関分析・信頼性分析

今回行った OSPAN, RSPAN, DSPAN (F), DSPAN (R) の結果を Table 1 に示した。平均得点について、OSPAN (P) は 69.03% ( $SD = 11.47$ ) で 95%CI は [66.42, 71.63], OSPAN (A) は 37.26% ( $SD = 15.62$ ) で 95%CI は [33.71, 40.80], RSPAN (P) は 70.32% ( $SD = 13.74$ ) で 95%CI は [67.20, 73.43], RSPAN (A) は 35.71% ( $SD = 18.33$ ) で 95%CI は [31.55, 39.87], DSPAN (F) は 56.82% ( $SD = 17.19$ ) で 95%CI は [52.92, 60.72], DSPAN (R) は 46.56% ( $SD = 18.31$ ) で 95%CI は [42.40, 50.72] となり、いずれも平均得点が 50% 付近であった。OSPAN では平均得点  $\pm 2SD$  が 0% - 100% に収まっており、小林・大久保 (2012) や大塚・宮谷 (2012) と異なり幅広い得点分布が得られた。

Table 1

尖度について、OSPAN (P) は.29, OSPAN (A) は.05, RSPAN (P) は-.48, RSPAN (A) は-.56, DSPAN (F) は-.20, DSPAN (R) は.30 であり、OSPAN (A) が最もピークが短く裾が長い分布となった。歪度については OSPAN (P) が -.36, OSPAN (A) は.41, RSPAN (P) は-.47, RSPAN (A) は.60, DSPAN (F) は.55, DSPAN (R) は.41 であり、OSPAN (P) が最も正負の歪みが少なく、次いで OSPAN (A) が少ない分布となった。正規分布は尖度 0, 歪度 0 のため、OSPAN は RSPAN に比べ正規分布に近い分布となった。

次に OSPAN, RSPAN, DSPAN (F), DSPAN (R) の結果について行った相関分析の結果を Table 2 に示した。

Table 2

全てのワーキングメモリ課題について互いに有意な正の相関があり ( $p < .05$ ), OSPAN (A) と RSPAN (A) ( $r = .40, p < .01$ ), OSPAN (A) と DSPAN (R) ( $r = .45, p < .01$ ), OSPAN (P) と DSPAN (R) ( $r = .48, p < .01$ ), DSPAN (F) と DSPAN (R) ( $r = .55, p < .01$ ) には有意な中程度の正の相関があった。

OSPAN の妥当性を検討するため, DSPAN (F) の得点を制御変数として, OSPAN と RSPAN の偏相関分析を行った。その結果, OSPAN (P) と RSPAN (P) に有意な正の相関が認められ ( $r_{\text{partial}} = .28, p < .05$ ), また OSPAN (A) と RSPAN (A) においても有意な正の相関が認められた ( $r_{\text{partial}} = .33, p < .01$ )。

各尺度の信頼性を検討するため, Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle (2004) にならいクロンバックの  $\alpha$  信頼性係数を算出した。信頼性係数は, 各尺度のスパンごとの得点を 1 項目あたりの得点として計算した。その結果, OSPAN (P) は.70, OSPAN (A) は.68, RSPAN (P) は.75, RSPAN (A) は.74, DSPAN (F) が.55, DSPAN (R) が.68 であった。

### 考察

本研究では日本人参加者向けに改良した OSPAN の妥当性について, 荳阪・荳阪 (1994) の日本語版 RSPAN や DSPAN との比較を行い検討した。まず OSPAN の計算課題正答率について, 本研究では 93% であった。これに対し Unsworth et al. (2005) では平均正答率 96% で 95%CI は [95.5, 96.5], また小林・大久保 (2012) では実験 1 の平均正答率は 96.6% で 95%CI は [95.82, 97.39], 実験 2 の平均正答率は 96% で 95%CI は [95.36, 96.80] であった。ここから本研究では計算課題の難易度が上昇したことが示唆された。その一方で計算問題の平均反応時間は 4208 ms であり, 95%CI は [3977, 4439] であった。Unsworth et al. (2005) では平均反応時間が 4155 ms



で 95% *CI* は [3958, 4352] であったことから, Unsworth et al. (2005) と比較して課題難易度が上昇しても計算にかかる時間が大きく増加しなかった。ここから, 計算にかかる時間を Unsworth et al. (2005) と同等にしつつ, 難易度を高めるという調整が適切になされたと考えられる。

OSPAN は RSPAN と比較し尖度・歪度が低かった。ここから, OSPAN は既存のワーキングメモリ尺度と同様に分布に偏りが無い尺度であることが示唆される。また Unsworth et al. (2005) の OSPAN を実施した小林・大久保 (2012) や大塚・宮谷 (2012) に対し, OSPAN (P)(A) は 95% 信頼区間と平均得点  $\pm 2SD$  がともに 0%100% の内に収まっており, ワーキングメモリ容量を幅広く測定できるものであると考えられる。

OSPAN と RSPAN には有意な正の相関があった。このことから本研究で行った OSPAN は, 荳阪・荳阪 (1994) の日本語版 RSPAN 等, 既存の尺度と同様にワーキングメモリ容量を測定できる妥当なワーキングメモリ尺度であると考えられる。また今回 OSPAN は OSPAN (P), OSPAN (A) とともに DSPAN (R) に中程度の相関があった。DSPAN (R) は数字列の保持だけでなく逆順にさせる処理が必要であり, DSPAN (F) よりもワーキングメモリにおける処理に関わる課題であるとされる (Morris & Jones, 1990; Psychological Corporation, 2002)。ここから OSPAN の成績にはワーキングメモリの保持だけでなく処理の要素も強く反映されている可能性が示唆される。

OSPAN の妥当性を検討するため, DSPAN (F) の得点を制御変数とし, OSPAN と RSPAN の偏相関係数を算出したところ, OSPAN と RSPAN に有意な正の相関があった。ここから OSPAN と RSPAN の変動は主に貯蔵機能を測定する DSPAN (F) の個人差を取り除いても残

ることがわかり、ワーキングメモリ尺度として妥当なものであることが示唆される。本研究では、Engle et al. (1999) が行ったようにレーヴン漸進マトリクスなど流動性知能を測定し、OSPAN との関連を検討する試みは行わなかった。ワーキングメモリの構成概念的妥当性の検討のため、また、ワーキングメモリが高次の心的処理に果たす役割を検討するため、今後は流動性知能などとの関連をみることも必要だろう。

各尺度の信頼性を検討するため OSPAN, RSPAN, DSPAN についてクロンバックの  $\alpha$  信頼性係数を算出したところ、OSPAN (P) は.70, OSPAN (A) は.68, RSPAN (P) は.75, RSPAN (A) は.74, DSPAN (F) が.55, DSPAN (R) が.68 であった。これに対し Kane et al. (2004) において OSPAN の  $\alpha$  信頼性係数は.80, RSPAN は.78, DSPAN は.79 であった (p. 199)。このように本研究の OSPAN と Kane et al. (2004) で実施した OSPAN で、信頼性係数の値に大きな違いはなかった。したがって本研究で開発した日本語版 OSPAN は先行研究と同様に信頼性のある尺度であると考えられる。

本研究では OSPAN を得点化する際、部分加点法と完全加点法を用いた。部分加点法は完全加点法に比べ検定力が高く、得点化の方法としてより適切であるとされる (e.g., Cowan, 2005; Friedman & Miyake, 2005)。また Unsworth et al. (2005) でも主な得点化方法として用いられていた (p.501)。そのため OSPAN は部分加点法による OSPAN (P) をワーキングメモリ容量の指標として用いるのが妥当であると考えられる。ただし OSPAN (P) の得点は OSPAN (A) と比較して高めになる傾向があるため、もし部分加点法による得点が満点付近に偏ってしまった場合は完全加点法を用いることも考えられるだろう。

Unsworth et al. (2005) では参加者に一般の市民も含ま

れ、年齢によらない幅広いワーキングメモリ容量の測定がなされていた。本研究における参加者は大学生・大学院生のみであったことから、今後は一般の市民を対象にしたり、ランダムサンプリングなどの方法を用いるなどして測定を行い、一般化可能性についてさらなる検討を進めることも必要だろう。ただし本研究の **OSPAN** は、大学生や大学院生を対象とした研究においては十分な妥当性・信頼性をもった尺度として実施できるものであると考えられる。

本研究で実施した **OSPAN** の特長として、再現性の高さ (Unsworth et al., 2005; Waters & Caplan, 1996) や、テスト – リテスト信頼性の高さ (Unsworth et al., 2005; Waters & Caplan, 1996)、数学不安が成績に影響を及ぼさない (Durette, 2011) ことに加え、コンピュータの使用により実験者の関与なしに測定できることで、実験者の測定への影響を排除できることがあげられる。**WAIS** における **DSPAN** や、荳阪・荳阪 (1994)、荳阪 (2002) の **RSPAN** では、刺激の提示などに実験者の関与が不可欠である。しかしながら実験者は全ての参加者に対し全く同じ測定を行うことは不可能であり、参加者への測定に影響を及ぼす可能性がある。これに対し本研究の **OSPAN** はすべてコンピュータ上で行うため、実験者や実験環境による影響がより少ないという利点のある測定方法であると考えられる。

ワーキングメモリにはさまざまな要素が関わっており (e.g., Baddeley & Hitch, 1974; Baddeley, 2000)、正確なワーキングメモリ容量測定のため **RSPAN** や **OSPAN** 等複数の尺度を用いる試みがなされている (e.g., Beilock & Carr, 2005; Tsai, Kang, & Peterson, 2010)。本研究で日本人向けに改良した **OSPAN** が開発されたことで、わが国におけるワーキングメモリ研究においても複数の尺度による精度の高い測定が可能になったと考えられる。

## 11. 参考文献

- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, **4**, 417-423.
- Baddeley, A. D. (2007). *Working Memory, Thought, and Action*. Oxford University Press.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8). New York: Academic Press. pp.47-89.
- Beilock, S. L. & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science*, **16**, 101-105.
- Case, R., Kurland, M. D., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, **33**, 386-404.
- Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. N. (Eds). (2008). *Variation in Working Memory*. UK: Oxford University Press.
- Cowan, N. (2005). *Working memory capacity*. Hove, East Sussex, UK: Psychology Press.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, **19**, 450-466.
- Durette, R. T. (2011). "Operation span task's suscep-

- tibility to math anxiety: Support from fluid intelligence” *UNLV Theses/Dissertations/Professional Papers/Capstones*. Paper 1002.
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, **11**, 19-23.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, **128**, 309-331.
- Friedman, N. P., & Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behavior Researching Methods*, **37**, 581-590.
- Jaeggi, S. M., Studer-Luechi, B., Buschkuhl, M., Su, Y., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning - implications for training and transfer. *Intelligence*, **38**, 625-635.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology: General*, **133**, 189-217.
- 小林晃洋・大久保街亜 (2012). 日本人参加者における作業記憶容量の測定 -- オペレーション・スパン・テストを用いて -- 専修人間科学論集 心理学篇 **2**, 27-34.

- (Kobayashi, A. & Okubo, M. (2012). The measurement of working memory capacity in Japanese participants using the Operation Span Test. *Bulletin of Senshu University School of Human Science: Psychology*, **2**, 27-34.)
- Morris, N., & Jones, D. M. (1990) . Memory updating in working memory: The role of the central executive. *British Journal of Psychology*, **81**, 111-121.
- National Council of Teachers of Mathematics. (2000). *Principles and Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.
- OECD. (2010). *PISA 2009 Results: What Students Know and Can Do -- Student Performance in Reading, Mathematics and Science (Volume I)*, <http://dx.doi.org/10.1787/9789264091450-en> (2012年10月15日参照)
- 荻阪満里子 (2002). 脳のメモ帳 ワーキングメモリ 新曜社
- 荻阪満里子・荻阪直行 (1994). 読みとワーキングメモリ容量 -- 日本語版リーディングスパンテストによる測定 -- 心理学研究 **65**, 339-345.
- (Osaka, M., & Osaka, N. (1994). Working memory capacity related to reading: Measurement with the Japanese version of reading span test. *Japanese Journal of Psychology*, **65**, 339-345.)
- 大塚一徳・宮谷真人 (2012). ワーキングメモリスパン課題における記銘項目数の影響, 日本認知心理学会第10回大会発表論文集 101.
- (Otsuka, K. & Miyatani, M.)
- Psychological Corporation. (2002). Updated WAISIII-WMS-III technical manual. San Antonio, TX: Psychological Corporation.
- Shelton, J. T., Elliott, E. M., Matthews, R. A., Hill, B. D.,

- & Gouvier, Wm. D. (2010). The relationships of working memory, secondary memory, and general fluid intelligence: Working memory is special. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **36**, 813-820.
- Stevenson, H. W., Chen, C., & Lee, S. (1993). Mathematics achievement of Chinese, Japanese, and American children: Ten years later. *Science*, **259**, 53-58.
- Stevenson, H. W., Lee, S., & Stigler, J. W. (1986). Mathematics achievement of Chinese, Japanese, and American children. *Science*, **231**, 693-699.
- Stevenson, H. W., Lee, S., Chen, C., Stigler, J. W., Hsu, C., & Kitamura, S. (1990). Contexts of Achievement: A Study of American, Chinese, and Japanese Children. *Monographs of the Society for Research in Child Development*, **55**, 1-190.
- Tsai, Y. D., Kang, S., & Peterson, M. S. (2010). Effects of individual differences on visual search task performance. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, **54**, 1981-1985.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, **28**, 127-154.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior research methods*, **37**, 498-505.
- Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, **17**, 635-654.
- Waters, G. S., & Caplan, D. (1996). The capacity theory of sentence comprehension: critique of Just and Carpenter (1992). *Psychological Review*, **103**, 761-772.
- 2012. 12. 16. 受稿, 2013. 9. 7. 受理 --

**12: Table 1: OSPAN, RSPAN, DSPAN の得点**

	平均値	<i>SD</i>	範囲	中央値	尖度	歪度	LQ	UQ	$\alpha$
OSPAN (P)	69.03	11.47	36.36 - 91.91	68.68	.28	-.36	66.42	71.63	.70
OSPAN (A)	37.26	15.62	4.04 - 80.81	36.36	.05	.41	27.27	45.45	.68
RSPAN (P)	70.32	13.74	40.00 - 91.43	71.43	-.48	-.47	62.86	81.43	.75
RSPAN (A)	35.71	18.33	5.71 - 75.71	32.86	-.56	.60	21.43	45.71	.74
DSPAN (F)	56.82	17.19	28.57 - 100.00	53.58	-.20	.55	44.05	66.67	.55
DSPAN (R)	46.56	18.31	7.14 - 100.00	44.05	.30	.41	35.71	53.58	.68
Unsworth et al. (2005), (P)	73.67	18.27	—	77.33	-1.14	1.41	64.33	90.67	—
Unsworth et al. (2005) (A)	52.21	23.21	—	50.00	-.02	-.49	37.33	68.00	—

A: 完全加点法, P: 部分加点法, LQ: 第 1 四分位点, UQ: 第 3 四分位点, 尖度・歪度・ $\alpha$  以外の単位は %



**13: Table 2: OSPAN, RSPAN, DSPAN の相関**

	OSPAN (P)	OSPAN (A)	RSPAN (P)	RSPAN (A)	DSPAN (F)
OSPAN (P)	—				
OSPAN (A)	.82**	—			
RSPAN (P)	.34**	.37**	—		
RSPAN (A)	.34**	.40**	.88**	—	
DSPAN (F)	.24*	.32**	.36**	.31**	—
DSPAN (R)	.48**	.45**	.38**	.28*	.55**

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

14: Figure 1

(6 - 3) × 5 = ?

答えがわかったら  
クリックしてください

**問題**

18

**回答**

P

**記録**

提示された文字を順番に答えてください  
忘れたら「?」を選んでください

1 F	H	J
K	L	3 N
2 P	4 Q	R
S	T	Y
?		

FPNQ

**フィードバック**

あなたは4文字中4文字思い出しました

現在の正答率は75%です

正答率が85%を下回っています

Figure 1. オペレーションスパンテストの試行例

