

事前テストにおける誤答と記憶定着

田中 紗枝子・宮谷 真人

学校現場で実施されるテストは、主に学習成果の評価を目的として使用されているが、テストを受けることそれ自体が、学習や長期的な記憶保持を促進することが知られている。また、近年、テストの解答が誤りであっても、後の評価テストの成績が向上することが明らかになってきている（プレテスト効果）。プレテスト効果を効果的な授業方法の開発につなげるためには、学習者の個人特性の影響を考慮する必要があるが、プレテスト効果による記憶促進が学習者の特性によって異なるかどうかについては、現在まで検討されていない。そこで本研究では、単語の対連合学習を用いた従来の研究方法を踏襲し、日本語の刺激を用いてプレテスト効果が出現するかどうかを確認した（実験1）。さらに、学習者の個人特性としてワーキングメモリ容量を測定し、その個人差がプレテスト効果に与える影響について検討した（実験2）。2つの実験の結果、テストを経験した群はそうでない群より評価テストの成績が高く、先行研究同様のプレテスト効果が確認された。しかし、プレテスト効果とワーキングメモリ容量の関連はなかった。このことから、プレテスト効果はワーキングメモリ容量の大小に依存しない現象である可能性が示された。

キーワード：記憶、プレテスト効果、対連合学習、誤答、ワーキングメモリ

Incorrect Answer in Pretest and Memory Fixation

Saeko Tanaka and Makoto Miyatani

In schools, tests are given mainly to evaluate learning achievements. Receiving a test is also known to promote learning and long-term retention of memory. Recently, even incorrect test answers have been revealed to improve scores in later evaluation tests (pre-test effect). To develop an effective teaching method by using the pre-test effect, it is necessary to understand how the effect differs between individual students depending on their characteristics. However, it has not been investigated whether the pre-test effect in promoting long-term retention of memory differs by the characteristics of students or not. In this study, the authors examined whether the pre-test effect appeared or not in a paired-associate learning task of new words, which has been widely used in studies on learning, by using Japanese stimulus words (Experiment 1). Then, the working memory capacity of each student was measured as an index for personal characteristics, and the effects of the personal difference on pre-test effect was investigated (Experiment 2). The experiments showed that the group that experienced pretest scored better in a subsequent evaluation test than the group that did not receive a pretest, confirming the pre-test effect as in preceding studies. However, no relationship was found between the pre-test effect and working memory capacity.

Therefore, the pre-test effect is possibly a phenomenon independent of working memory capacity.

Key Words: Memory, Pre-test Effect, Paired-associate Learning, Incorrect Answer, Working Memory

問題と目的

学校現場において、テストは知識や技能がどれだけ定着しているかを評価するために実施される。また記憶研究においても、再生テストや再認テストがパフォーマンス評価の指標として利用されている。しかし、能力評価のためのツールという側面以外にも、記銘すべき材料に関するテストを事前に受けることが、記銘材料を複数回学習することと同程度に学習にとって効果的であることが知られている。この現象は“テスト効果”と呼ばれ、多くの研究が行われている（遠藤，2007; Richland & Karpicke, 2006; 多鹿，2008）。

テストにおける正答・誤答が記憶定着に及ぼす影響

では、テストにおける正答と誤答では、どちらがより効果的なのだろうか。これまでは、行動主義における学習理論などの立場から、テストにおける誤答は正答（評価テスト時に解答すべき情報）と競合を起し、その結果記憶を抑制してしまうため、事前テストでの正答時にのみ記憶の促進効果があるといわれてきた。しかし近年、テスト時に誤った解答をする、つまり誤情報を想起することが、正しい情報の学習を促進させることが分かってきた。Kornell, Hays, & Bjork (2009) や Richland, Kornell, & Kao (2009) は、意味的に弱い関連のある2つの単語（手がかり単語と評価テスト時に想起するターゲット単語, e.g. “tide” と “beach”）の対連合学習課題を用いた実験を行い、誤情報想起による“プレテスト効果”を報告している。この研究では、実験参加者は Pretest 群と No-Pretest 群のいずれかに分けられ、対連合学習の後、手がかり単語をもとにターゲット単語を再生する評価テストを行った。その際、Pretest 群は単語対の学習の前に、手がかり単語から連想される単語（e.g. wave）を1つ生成する事前テストを行った。その結果、Pretest 群は No-Pretest 群よりも評

価テストにおける再生成績が有意に高かった。Kornell et al. (2009) 等の研究は刺激として単語対を使用したものであるが、同様の結果は散文を読みその内容についてのテストを行った研究（Richland et al., 2009）や、大学の講義内容に関する事前テストを行った群と行わなかった群の評価テスト成績について比較した研究（Butler & Roediger, 2007）でも得られている。

これらのプレテスト効果が生起するメカニズムについては、これまでに一貫した知見は得られていないが、連想語が手がかり語とターゲット語の媒介（mediator）として機能するという説と、精緻化を促すという説が提案されている。媒介説によると、プレテストによる記憶促進は、手がかり語から連想語を想起し、その連想語をもとにターゲット語を再生することができるため生起すると考えられる（Pyc & Rawson, 2010）。一方の精緻化説によれば、手がかり語からの連想を行うと、連想語やターゲット語を含む意味のまとまり（意味ネットワーク）が活性化され、それが手がかりとなってターゲット語の記憶が促進されると考えられている（Grimaldi & Karpicke, 2012; Hays, Kornell & Bjork, 2013）。

プレテスト効果に影響を及ぼす個人差要因

どの人にとっても、つまり個人の特性によらずプレテスト効果は出現するのだろうか。誤情報の検索が正情報の記憶を促進するというプレテスト効果について、学習者の個人差の影響について直接検討した研究は見当たらないが、それに関連するものとして、以下のような報告がある。

モニタリング能力 Grimaldi & Karpicke (2012) や Hays et al. (2013) の指摘するとおり、多くの情報が含まれる意味ネットワークから、想起すべき正情報のみを引き出すことによって記憶が促進されているのであれば、誤情報か正情報かの判断を行う能力の個人差

によってプレテスト効果の表れ方に違いがあると考えられる。このような判断を行うための能力としてソースモニタリング能力が挙げられており、この能力はワーキングメモリによって支えられているといわれている。ワーキングメモリとは、複雑な認知活動に欠かせない、情報の処理と保持を同時に行う認知システムであるが、その容量には制限がある。このワーキングメモリの容量はスパンテストによって測定されるが、その結果は言語処理や推論、問題解決などの高次認知活動の成績と関係があることから、ワーキングメモリはこれらの認知機能を支える機構であると考えられている (e.g. 苧阪, 2006; 土田, 2009)。

虚記憶に関する研究では、ワーキングメモリ容量とソースモニタリング能力の関連について、ワーキングメモリ容量の大きい人ほどソースモニタリング能力も優れており、虚再生や虚再認が起きにくいことが、様々なパラダイムを用いた研究によって示されている。Leding (2012) によれば、これは、大きなワーキングメモリ容量を持つ人は、自分の記憶の情報源に関するモニタリングを低容量の人に比べてより正確に行うことができるためであり、その結果虚記憶を引き起こすような検索過程においても、正しい学習の判断ができる。

この考えにもとづけば、意味ネットワークに含まれている情報が解答すべき正情報なのか、それともそれ以外の情報なのかという判断を必要とする課題で観察されるプレテスト効果についても、虚記憶研究と同様の結果が得られると予測される。つまり、ソースモニタリング能力に優れたワーキングメモリ容量の大きい参加者ほど、事前テストによる記憶促進効果が大きくなると考えられる。

方略の利用 一方、ワーキングメモリ容量を測定するスパンテストの成績について、課題遂行の際に使用する方略の個人差が、スパンテスト成績に影響を与えることが知られている (遠藤・苧阪, 2012; 齊藤・三宅, 2000)。

遠藤・苧阪 (2012) は、リーディングスパンテスト遂行時に利用する方略について、容量低群と高群の比較を行っている。その結果、容量高群は容量低群に比べて効率的な方略をより多く使用していた。このことから、ワーキングメモリ容量の小さい参加者は記憶のために効果的な方略を選択できていないという可能性が考えられる。実際、効果的な方略を利用するよう訓練を行うことで、スパンテストの得点が上昇するという結果も得られている (McNamara & Scott, 2001)。

この結果にもとづけば、ワーキングメモリ容量が小さい参加者にとっては、記憶の為の方略を生成しなければならない No-Pretest 条件よりも、誤情報を生成するという方略を与えられる Pretest 条件においてより大きな記憶の促進効果があると考えられる。一方、ワーキングメモリ容量が大きい人は個人が持つ効果的な方略を使用するため、No-Pretest 条件でも高い成績を得ると考えられる。その結果、No-Pretest 条件と Pretest 条件の成績差は容量低群の方が大きくなると考えられる。

本研究の目的

以上を踏まえ、本研究ではワーキングメモリ容量の個人差がプレテスト効果に及ぼす影響について検討する。もし、ワーキングメモリ容量高群が容量低群に比べて事前テストによる記憶促進効果が大きければ、プレテスト効果が生じるメカニズムとして、事前テスト時に得られた情報の正誤を区別するためのソースモニタリングを想定できる。また、事前テストは、ソースモニタリング能力が低いといわれている容量低群にとっては効果的な記憶促進法ではないことが示唆される。一方、No-Pretest 条件と Pretest 条件の成績差がワーキングメモリ容量高群より低群で大きければ、事前テストという記憶のための方略を提示したことにより、容量低群の成績がより向上したと考えられ、事前テストは、容量低群にと

ってより適切な方略であるといえるだろう。また、もしワーキングメモリ容量の高低にかかわらず事前テストを行ったことによる記憶の促進効果が得られた場合、事前テストは個人差を考慮する必要のない記憶方略であるといえることができるだろう。いずれにしても、プレテスト効果に影響を及ぼす個人差要因について検討することは、新たな授業方法を開発していく上で重要であると考えられる。

そこでまず実験 1 として、日本語による刺激を作成して Kornell et al. (2009) と同様の実験を行い、以降の実験でプレテスト効果について検討するのに適切な刺激語リスト等を確定する。続く実験 2 では、単語の対連合学習の前にワーキングメモリ容量を測定し、それにもとづいて参加者をワーキングメモリ容量高群と低群に分け、容量の違いによってプレテスト効果の現れ方に違いがあるかどうかを検討する。

日本語刺激を用いたプレテスト効果の検討 (実験 1)

プレテスト効果について検討するための日本語の刺激を選定することを目的として実験 1 を行った。

方法

実験参加者 大学生、および大学院生 24 名（うち男性 9 名、女性 15 名）が実験に参加した。年齢は 19—25 歳（平均 20.13 歳、SD=1.51）であった。このうち、11 名（うち男性 4 名）を Pretest 群、13 名（うち男性 5 名）を No-Pretest 群にランダムに振り分けた。参加者には実験の内容を口頭で説明した後、書面による同意を得た。

刺激 Kornell et al. (2009) の研究では、連想強度（ある手がかりとなる単語と、そこから連想される“ある単語”について、手がかり単語が提示されたときに、その単語を連想する人の割合）をもとに、手がかり単語とターゲット単語を選択していた。そこで本実験

では、同様のものを調査した水野（2011）を参考に、連想強度が .045—0.054 となるような手がかり単語とターゲット単語の単語対 60 組（計 120 単語）を選出した。手がかり単語は 3 モーラ、ターゲット単語は 2—4 モーラであり、単語の表記には漢字、ひらがな、カタカナが含まれていた。

課題 課題は、Kornell et al. (2009) と同様の学習フェーズ、挿入課題、評価テストの 3 つで構成した。なお、先行研究では事前テストと評価テスト時の単語解答をキーボードからの入力によって行っていたが、日本語では変換作業のための余分な時間が必要となるため、解答は口頭で行わせ、連想語の視覚的なフィードバックは与えなかった。なお、実験中呈示する画面は先行研究と同じであった。

学習フェーズでは、Pretest 群は事前テストと手がかり単語とターゲット単語の対連合学習を、No-Pretest 群は手がかり単語とターゲット単語の対連合学習のみを行った。Pretest 群の学習フェーズでは、まずパソコンのディスプレイ上に手がかり単語と枠が 7 秒間呈示され、参加者はその手がかり単語からターゲット単語になるような単語（以下“生成単語”）を連想して、口頭で解答した（事前テスト）。7 秒経つと枠が消え、その 500 ミリ秒後にさらに画面が切り替わり、正しいターゲット単語が手がかり単語とともに 5 秒間表示された。参加者は 5 秒間の間に、先ほどの生成単語ではなく、ターゲット単語を手がかり単語と対にして覚えるよう指示された（対連合学習）。500 ミリ秒の後、また次の手がかり単語に対する事前テストと対連合学習を行った。一方の No-Pretest 群の学習フェーズでは、Pretest 群と同じ内容の対連合学習のみを行った。なお、単語対の呈示順序は参加者ごとにランダムであった。

挿入課題では、暗算課題（2—3 桁の整数の四則計算）を 5 分間行った。パソコンのディスプレイ上に数式と枠が表示され、参加者は

テンキーを用いて解答を画面上の枠の中に入力した。

評価テストでは、両群とも同一の手がかり再生課題を行った。ディスプレイ上に手がかり単語と枠が7秒間表示され、参加者はその間に手がかり単語と対になっていたターゲット単語を思い出し、口頭で解答した。7秒後、次の手がかり単語に対する再生を行った。なお、評価テストにおいても手がかり単語の表示順序は参加者ごとにランダムであった。

手続き まず全ての参加者に、これから単語対の学習を行い、後でどれだけ覚えているかテストを行うことを説明した。単語学習課題の前に、練習を10題行った。ここでの成績は分析では使用しなかった。その後、Pretest群は事前テストと対連合学習からなる学習フェーズを、No-Pretest群は対連合学習のみの学習フェーズを行った。単語対60組の学習がすべて終わったのち、5分間の挿入課題を行った。参加者はできるだけ速く正確に、5分間問題を解き続けるよう指示された。挿入課題を行った後、評価テストを行った。学習フェーズで使用した60組の単語対すべての再生が終わったところで実験を終了した。実験の所要時間は、Pretest群が35—40分程度、No-Pretest群が25—30分程度であった。

結果

Pretest群において、事前テスト時にターゲット単語を連想した（生成単語とターゲット単語が同一であった）割合は6.0%であった。ターゲット単語を連想した手がかり単語については、分析から除外した。

プレテスト効果 評価テストでの再生成績に対する事前テストの効果について、Pretest群とNo-Pretest群の間で平均値に差が見られるかどうか、対応のないt検定を行った（Figure 1）。再生率はそれぞれ、Pretest群が.88（SD = 0.11）、No-Pretest群が.72（SD = 0.18）であり、Pretest群の方がNo-Pretest群

より有意に成績が高かった（ $t(20.18) = 2.64$, $p < .05$ ）。

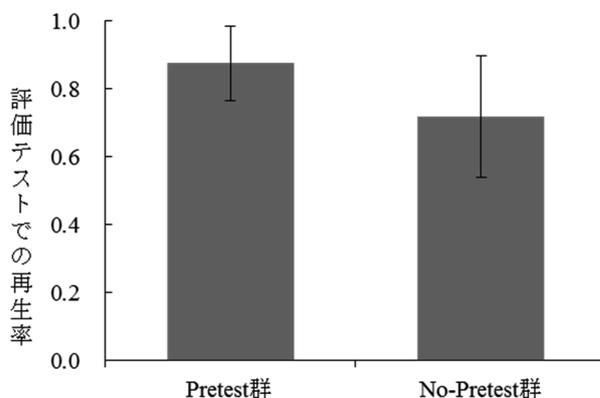


Figure 1. 実験1における評価テストでの再生率（誤差線は標準偏差）

挿入課題の成績とプレテスト効果の関連 プレテスト効果に影響を及ぼす個人差要因について検討するため、挿入課題の成績と評価テストでの再生成績の関連について、群別にピアソンの積率相関係数を算出した。その結果、No-Pretest群の挿入課題の正答率と再生成績の間に中程度の正の相関があった（ $r = .49$, Figure 2）。

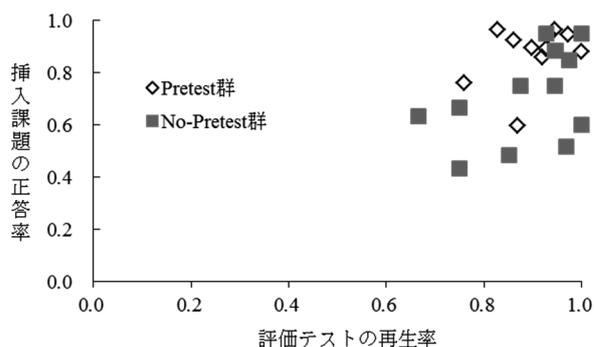


Figure 2. 挿入課題の正答率と評価テストの再生率の相関

そこで、学習フェーズ時の群ごとに正答率が中央値より高い者を成績高群、低い者を成績低群とし（ちょうど正答率が中央値に当たる参加者は含まなかった）、評価テストにおける再生成績（Table 1）について、学習フェー

ズ時の群（2：Pretest 群，No-Pretest 群）×挿入課題の成績（2：高群，低群）の2要因分散分析を行った。その結果，学習フェーズ時の群の主効果のみ有意であり（ $F(1, 17) = 6.28, p < .05$ ），挿入課題の成績の主効果と交互作用は有意ではなかった（それぞれ $F(1, 17) = 2.28, n.s.$; $F(1, 17) = 0.01, n.s.$ ）。

Table 1
挿入課題の成績別の評価テスト再生率

		参加者数	評価テスト再生率	
			平均	SD
Pretest群	正答率 高群	3	.93	.04
	低群	7	.76	.19
No-Pretest群	正答率 高群	4	.83	.15
	低群	6	.65	.19

考察

日本語刺激によるプレテスト効果 実験1の目的は，Kornell et al. (2009) 等の結果が日本語の刺激を用いても再現されるかどうかを確認し，今後の研究で用いる刺激を確定することであった。まず，学習フェーズの群ごとに成績を比較した結果，Pretest 群のほうが，No-Pretest 群より評価テストにおける再生成績が高く，先行研究同様，事前テストを行ったことにより記憶が促進されていた。実験1で用いた日本語の刺激は，プレテスト効果を検討するのに適切な刺激であったといえるだろう。

挿入課題の成績とプレテスト効果 No-Pretest 群において，評価テストの再生成績と，挿入課題（暗算）の成績の間に中程度の正の相関があり，No-Pretest 群において挿入課題の成績がよい人ほど，再生成績もよいことがわかった。暗算は，その遂行にワーキングメモリ容量が深くかかわっているといわれている（e.g. 齊藤・三宅，2000）。この知見にもとづくと，実験1における挿入課題である暗算の成績が高かった，つまりワーキングメモリ容量の大きい参加者ほど，No-Pretest 条件で

あっても成績が高いといえる。ワーキングメモリ容量の小さい参加者は，事前テストがない場合にはワーキングメモリ容量の大きい参加者に比べて成績が低くなるが，事前テストを行うことによって，容量の大きい参加者と同程度の再生成績を上げることができたといえるだろう。

No-Pretest 群で成績の良かった参加者に，単語対を記憶する際に使用した方略について尋ねたところ，手がかり単語とターゲット単語が似た意味である，手がかり単語をカタカナ化したものであるなどの基準にもとづいて，呈示された単語対をグループ化し，評価テスト時の手がかりとしていたことが分かった。このことは遠藤・苧阪（2012）で示されているとおり，ワーキングメモリ容量の大きい人ほど，記憶に効果的な方略を自ら選択し，使用することができていると考えることができる。一方のワーキングメモリ容量が小さい参加者は，方略の指定されない No-Pretest 条件では成績が低くなるが，事前テストという形で記憶する際の方略が示されることにより，容量の大きい参加者と同程度にまで成績が向上したと考えられる。つまり，プレテスト効果がワーキングメモリ容量の小さい人にとって，より有効である可能性が示唆されたといえるだろう。

ワーキングメモリ容量がプレテスト効果に及ぼす影響（実験2）

目的

実験1の結果から，高ワーキングメモリ容量の人は，事前テストの有無（学習時の方略の呈示）によって成績は変わらないが，低ワーキングメモリ容量の人は，学習時に記憶のための方略として事前テストを行うことで，より再生成績が高くなると考えられる。そこで実験2では，ワーキングメモリ容量の違いによって事前テストが単語の再生成績に及ぼす効果が異なるかどうかを検討する。

ワーキングメモリ容量の測定には、Daneman & Carpenter (1980) によって開発されたリーディングスパンテストが最も一般的に使用されている (齊藤・三宅, 2000)。しかし、より正確にワーキングメモリ容量を測定する際には、複数のスパンテストを組み合わせる方が望ましいとされている (e.g. Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005; 遠藤・苧阪, 2012; 大塚・宮谷, 2007)。そこで実験 2 ではリーディングスパンテストの他にオペレーションスパンテスト (Turner & Engle, 1989) を使用して、ワーキングメモリ容量の測定を行った。

方法

参加者 2 府県の計 4 つの大学に在籍する大学生、および大学院生 71 名 (うち男性 17 名、女性 54 名) が実験に参加した。年齢は 18—25 歳 (平均 21.8 歳, SD=1.61) であった。このうち、36 名 (うち男性 9 名) を Pretest 群、35 名 (うち男性 8 名) を No-Pretest 群にランダムに振り分けた。参加者には実験の内容を口頭で説明した後、書面による同意を得た。

課題 実験 2 では、2 種類のワーキングメモリ測定課題 (スパンテスト) と、実験 1 と同じ単語学習課題の合計 3 つの課題を行った。スパンテストとして、Unsworth, Heintz, Schrock, & Engle (2005) が作成した自動オペレーションスパンテストとリーディングスパンテストを日本語化したものを使用した。

単語学習課題は実験 1 と同じものであり、Pretest 群は学習フェーズとして事前テストと対連合学習を、No-Pretest 群は対連合学習のみを行い、その後挿入課題、評価テストを行った。

手続き はじめにスパンテスト 2 種類を行い、その後単語学習課題の順で課題を実施した。このうち、オペレーションスパンテストとリーディングスパンテストの順序は参加者ごとにランダムとした。2 種類のスパンテストを行った後で、単語学習課題を行った。単

語学習課題は、実験 1 と同様に学習フェーズ、挿入課題、評価テストの順に実施した。なお、本実験の所要時間は、Pretest 群が 70 分程度、No-Pretest 群が 60 分程度であった。

結果

スパンテストの成績 スパンテストの得点化については、一般的に Daneman & Carpenter (1980) が提案する“伝統的なスパン得点”が使用されることが多い。しかし近年、伝統的なスパン得点は最後のアイテムサイズが得点となるという性質上、参加者の能力が課題の難易度と強い関連を持ってしまふことが指摘されている (Conway et al., 2005)。そこで本研究では、複数の研究において妥当性が示されている正再生率 (その試行中、正しく解答できたアイテムの割合の平均値) をスパンテストごとに算出した。各スパンテストの成績の平均値は、リーディングスパンテストが .78 (SD = 0.12)、オペレーションスパンテストが .80 (SD = 0.13) であり、スパンテスト間の相関係数は $r = .64$ ($p < .01$) であった。また、これらを平均したものを“複合スパン成績”とした。スパンテストの成績と挿入課題の成績 (解答数、正答数、正答率) との間に関連がみられるかどうか分析を行ったところ、相関係数は有意ではなかった (Table 2)。以降の分析では、この複合スパン成績を、個人のワーキングメモリ成績として使用した。

Table 2

スパンテストの成績と挿入課題の成績との相関

	解答数	正答数	正答率
リーディングスパン	-.01	.04	.17
オペレーションスパン	.05	.09	.16
複合スパン	.02	.08	.18

プレテスト効果 Pretest 群において、事前テスト時にターゲット単語を連想した割合は 5.2% であった。以降は、これらの単語を除外して分析を行った。

評価テストでの再生成績に対するプレテスト効果について、Pretest 群と No-Pretest 群の間で平均値に差が見られるかどうか、対応のない t 検定を行った。その結果、Pretest 群の再生率が .81 (SD=0.09)、No-Pretest 群の再生率は .72 (SD=0.21) であり、Pretest 群の方が有意に成績が高く ($t(44.63) = 2.37, p < .05$)、実験 2 においても有意なプレテスト効果が出現した。

プレテスト効果と挿入課題の成績について、実験 1 と同様に相関係数を算出した。その結果、本実験では Pretest 群においては弱い正の相関が、No-Pretest 群においては弱い負の相関が（特に正答数との間に）観察された (Table 3)。

Table 3
評価テストの成績と挿入課題の成績との相関

	解答数	正答数	正答率
Pretest群	.28 *	.32 *	.16
No-Pretest群	-.30 *	-.30 *	-.08

* $p < .10$

プレテスト効果に対するワーキングメモリ容量の影響 プレテスト効果に対するワーキングメモリ容量の影響を検討するため、複合スパンテスト成績が中央値より高い参加者を上位群、低い参加者を下位群とした。なお、複合スパン成績の中央値は .82 であり、成績が中央値と等しかった参加者は分析から除外した。

ワーキングメモリ容量によってプレテスト効果の現れ方に違いがあるかどうか、学習フェーズ時の群 (2 : Pretest 群, No-Pretest 群) × スパン成績 (2 : 高群, 低群) の 2 要因分散分析を行った。その結果、有意であったのは学習フェーズ時の群の主効果のみであり ($F(1, 63) = 4.43, p < .05$)、Pretest 群の方が No-Pretest 群より再生成績は高かった。スパン成績の主効果、および交互作用は有意ではなかった (それぞれ $F(1, 63) = .38, n.s.$; $F(1, 63)$

$= .04, n.s.$)。各群の参加者数、および平均値と SD を Table 4 に示した。

Table 4
学習フェーズ時の群と複合スパン成績別の評価テスト再生率

	参加者数	評価テスト再生率		
		平均	SD	
Pretest群	スパン 高群	14	.79	.10
	低群	20	.83	.08
No-Pretest群	スパン 高群	19	.71	.20
	低群	14	.73	.25

考察

実験 2 の目的は、ワーキングメモリ容量の大小によって、事前テストが単語の再生成績に及ぼす効果の大きさが異なるという可能性について検討することであった。

ワーキングメモリ 実験 2 で測定したリーディングスパン得点とオペレーションスパン得点の間には高い相関がみられ、このことから 2 つのスパンテストは単一のものを測定していると考えられる。しかし、挿入課題として行った暗算の成績との間にはほとんど相関がなく、先行研究 (e.g. Hecht, 2002) とは異なる結果となった。その理由として、本実験の暗算課題が、記憶課題の途中で行われた挿入課題であったことが挙げられる。つまり、参加者は後の評価テストのために資源をセーブしながら暗算課題に取り組んだため、相関が現れにくくなったと考えられる。また、実験 2 で用いたスパンテストが、暗算課題にはあまり関与しないワーキングメモリの側面を測定していた可能性もある。様々なスパンテストがどのような心的過程を反映しているのかについてはいまだに多くの議論や解釈が存在する (齊藤・三宅, 2000)。本実験で使用したスパン課題がプレテスト効果の個人差を検討するのに適切であったかどうかについて、さらに検討する必要があるだろう。

プレテスト効果 実験 2 においても No-

Pretest 群に比べて Pretest 群の再生成績が高いというプレテスト効果が得られた。一方、挿入課題の成績と群ごとの再生成績の関連については、実験 1 とは異なり、Pretest 群では弱い正の相関が、No-Pretest 群では弱い負の相関が観察された。つまり、ワーキングメモリ容量の大きい人ほど事前テストが効果的であり、ワーキングメモリ容量の小さい人は容量の大きい人ほど学習フェーズ時の群による成績の違いがないという結果であった。実験 1 と結果が異なった理由については、総合考察で検討する。

プレテスト効果に対するワーキングメモリ容量の影響 ワーキングメモリ容量の大小によって参加者を群分けし、プレテスト効果について検討したところ、有意な交互作用は得られず、ワーキングメモリ容量によるプレテスト効果の大きさの違いはなかった。Leding (2012) や遠藤・苧阪 (2012) の結果を踏まえれば、プレテスト効果にワーキングメモリ容量が関係しないとは考えにくいだが、本実験においてこのような結果が得られた理由として次のような可能性が考えられる。

まず 1 つは、本実験で利用したスパンテストが、プレテスト効果に影響を及ぼすようなワーキングメモリの側面を測定できていなかった可能性である。上述したとおり、本実験で使用したスパン課題と暗算課題はワーキングメモリの異なる側面と関連していたと考えられる。同様に、プレテスト効果に影響を及ぼすようなワーキングメモリの側面と、スパン課題によって測定した側面が異なっていたのではないか。しかし、暗算課題との相関がみられなかった理由としては、上記のとおり暗算を挿入課題として行ったためという可能性も考えられる。本実験で使用したスパン課題だけではなく、他の課題も用いて、プレテスト効果との関連を検討する必要があるだろう。

2 つ目は、プレテスト効果はソースモニタ

リングや記銘時の方略呈示などによるものではないため、ワーキングメモリ容量の影響を受けないという可能性である。Bixter & Daniel (2013) は虚記憶とワーキングメモリの関連について検討し、呈示した単語リストが虚記憶を生成するようなものになっていることを参加者に伝えなかった場合、ワーキングメモリと単語再生成績の間に関連がないことを報告している。これは、“虚記憶が生成される”という単語リストの性質に気付かせなければ、ソースモニタリング等へ配分される注意が小さくなるためであるといわれている。本実験でも、事前テスト時の誤情報を、評価テストの正情報再生時にどのように利用するかについては参加者に委ねられており、そのため参加者は評価テスト時の手がかりとして情報の正誤のソースモニタリングを利用していなかったのではないだろうか。ソースモニタリングに対して注意が分配されていなかったため、本研究でもワーキングメモリ容量がプレテスト効果に関連しなかったと考えられる。このことから、上述したようなモニタリング能力や方略の利用能力など、ワーキングメモリ容量が関わっていると考えられている能力の個人差はプレテスト効果に影響を及ぼさず、事前テストによる記憶の促進効果は、個人差によらない方略であるということができよう。しかし、ソースモニタリングが評価テスト時に利用されていないとすれば、Grimaldi & Karpicke (2012) の提案する意味ネットワークの活性化によってプレテスト効果が生じているとは考えにくく、誤情報が評価テスト時にどのように利用されているのかさらに検討する必要があるだろう。

総合考察

本研究の目的は、日本語の刺激を用いてプレテスト効果の存在を確認し、さらに、プレテスト効果に影響を与える個人差影響について、ワーキングメモリの観点から検討するこ

とであった。その結果、プレテスト効果とワーキングメモリ容量の間に関連はなく、プレテスト効果は個人差に依存しない現象である可能性が示された。このことから、例えば授業場面で事前テストを利用した Butler & Roediger (2007) のように、従来確認のためのテストを行うことで成績が低下するといわれてきたワーキングメモリ容量の小さい生徒に対しても、学習を定着させる方法として用いることができると考えられる。しかし、学校現場で使用するためには、どのようなテストがより効率的なのかなど、プレテスト効果が起こるメカニズムに関するさらなる研究が必要だろう。

また、挿入課題として行った暗算の成績と再生成績の関連は、実験間で一貫していなかった。実験 1 では暗算成績の低い参加者において、事前テストという記憶のための方略を示したことで成績が向上することが示唆された。一方実験 2 では、事前テストによって暗算成績高群の再生成績が向上し、プレテスト効果においてソースモニタリングが重要である可能性が示された。このような不一致を解消するためには、プレテスト効果を検討する実験パラダイムにおいて、事前テスト時の誤情報がどのように利用されているのかを詳しく検討する必要があるだろう。さらに、ワーキングメモリ容量がプレテスト効果を促進する働きと抑制する働きの両方を持つ可能性も考慮する必要があると思われる。

テスト効果におけるワーキングメモリの個人差の影響について検討した Brewer & Unsworth (2012) では、プレテスト効果ではなく、学習後に実施したテストが評価テストの成績に及ぼす影響(テスト効果)、およびワーキングメモリや注意機能などの認知能力の個人差がテスト効果に及ぼす影響について検討している。この研究においてもワーキングメモリの個人差によるテスト効果への影響はなく、エピソード記憶能力と一般流動性知能

の間に弱い負の相関がみられた。このことから、認知機能の個人差はテスト効果に直接は影響しないが、“より効果的な方略を使用できるようになる”という点において、認知能力の低い人にとって間接的な効果があると示唆されている。本研究の実験 1 において、挿入課題の成績低群でプレテストの効果がより大きく現れ、実験 2 においてはワーキングメモリ容量とプレテスト効果に関連がなかったことから、Brewer & Unsworth (2012) が検討したテスト効果と同様、事前テストはワーキングメモリ容量によらず効果的であるが、容量の小さい人にとっては、学習時の方略が提示されたことにより特に記憶が促進されたと考えられるだろう。

引用文献

- Bixter, M. T., & Daniel, F. (2013) . Working memory differences in illusory recollection of critical lures. *Memory & Cognition*, 41, 716-725.
- Brewer, G. A., & Unsworth, N. (2012) . Individual differences in the effects of retrieval from long-term memory. *Journal of Memory and Language*, 66, 407-415.
- Butler, A. C., & Roediger, H. L. (2007) . Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, 19, 514-527.
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005) . Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12, 769-786.
- Daneman, M., & Carpenter, P. A. (1980) . Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, 19, 450-466.
- 遠藤香織・苧阪満里子(2012) . 日本語版リーディングスパンテストにおける方略利用の

- 個人差 心理学研究, 82, 554-559.
- 遠藤正雄 (2007). 再生課題によるテスト効果
近畿福祉大学紀要, 8, 37-41.
- Grimaldi, P. J., & Karpicke, J. D. (2012). When and Why do retrieval attempts enhance subsequent encoding? *Memory & Cognition*, 4, 505-513.
- Hays, M. J., Kornell, N., & Bjork, R. A. (2013). When and why a failed test potentiates the effectiveness of subsequent study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 290-296.
- Hecht, S. A. (2002). Counting on working memory in simple arithmetic when counting is used for problem solving. *Memory & Cognition*, 30, 447-455.
- Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 989-998.
- Leding, J. K. (2012). Working memory predicts the rejection of false memories. *Memory & Cognition*, 20, 217-223.
- McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. *Memory & Cognition*, 29, 10-17.
- 水野りか (編) (2011). 連想語頻度表—3 モーラの漢字・ひらがな・カタカナ表記語—ナカニシヤ出版
- 苧阪満里子 (2006). ワーキングメモリ (1) 入門講座 臨床脳波, 48, 691-696.
- 大塚一徳・宮谷真人 (2007). 日本語リーディングスパン・テストにおけるターゲット語と刺激文の検討 広島大学心理学研究, 7, 19-32.
- Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, 330, 335.
- Richland, L. E., III, & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1, 181-210.
- Richland, L. E., III, Kornell, N., & Kao, L. S. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied*, 15, 20-27.
- 齊藤 智・三宅 晶 (2000). リーディングスパン・テストをめぐる6つの仮説の比較検討 心理学評論, 43, 387-410.
- 多鹿秀継 (2008). テストが学習材料の長期の記憶成績に及ぼす影響 親和女子大学大学院研究紀要, 4, 57-65.
- 土田幸男 (2009). ワーキングメモリ容量とは何か? —個人差と認知パフォーマンスへの影響— 北海道大学大学院教育学研究院紀要, 109, 81-92.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28, 127-154.
- Unsworth, N., Heintz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37, 498-505.

著者

田中 紗枝子 広島大学大学院教育学研究科
博士課程後期

宮谷 真人 広島大学大学院教育学研究科

本論文は、Theory and Research for Developing Learning Systems, Vol.1 所収の英語論文“Incorrect answer in pretest and memory fixation”の日本語訳論文である。