

学位論文

誤情報の検索が正情報の学習を促進するメカニズム
—検索する情報量とその処理の観点からの検討—

広島大学大学院教育学研究科

教育人間科学専攻

D144287

田中 紗枝子

目 次

第 1 章 本研究の背景と目的	
第 1 節 テストによる学習促進	2
第 2 節 誤検索効果のメカニズム	10
第 3 節 誤検索効果が意味ネットワークの活性化によるものである 根拠—遅延フィードバックによる影響—	16
第 4 節 本研究の目的	19
第 2 章 誤情報の検索が正情報の学習を促進するメカニズムの検討	
第 1 節 検索する情報量の影響（研究 1）	25
第 2 節 検索情報の解答方法の影響（研究 2）	48
第 3 節 注意分割の影響（研究 3）	52
第 3 章 総合考察	
第 1 節 本研究の成果と意義	63
第 2 節 今後の課題	71
引用文献	75
Appendix.....	82

第 1 章 本研究の背景と目的

第 1 節 テストによる学習促進

1. 学校教育の場におけるテスト

我々は学校教育の場において頻繁にテストを受ける。学校内で実施される一般的なものとしては、例えば 1 回もしくはもう少し長いスパンの授業で学んだ内容を確認するためのテスト、学期末などに実施される成績評価のためのテスト、卒業や進級・進学のためのテストがあるであろう。このほかにも、高校や大学など新たな学校への入学のためのテストやその模擬試験を受ける場合もある。さらに 2007 年からは全国の小学 6 年生と中学 3 年生全員を対象とした全国学力・学習状況調査も実施されているし、TOEIC や漢字検定などの資格試験を学校内で受けることもある。

中野（2007）によれば、テストを実施する目的は以下の 4 点にまとめられる。第一に子どもたちがいかによく学習しているか、いかに良く学習したかを知る。第二に学校を越えた広い範囲の子どもたちに（特に標準化された）テストを実施することによって、教室や学校の中だけでは得られないデータ（全体における個々の子どもの位置や学校の位置、あるいは過去のデータとの比較など）が得られる。第三にテストによって得られたデータに基づいて、教師や学校は自分の実践を振り返り、明日の実践を行う。そして第四にテストの結果に基づいて、適切な行政施策が行われる。このように、学校教育の場におけるテストは学習者の知識や能力を測ったり、教授者の実践を改善したりするために利用され、また教授者もテストを測定のためのデバイスとみなしている場合が多い（Roediger & Karpicke, 2006b）。しかしテストは知識や能力を測定するだけではなく、それを受けることによって直接的、もしくは間接的に学習

者の知識や能力が変化することが知られている。

2. テストが学習に及ぼす影響

2.1 直接効果

テストによる学習促進の直接効果とは、「テストを受けるという行為そのものが、他のプロセスを介さずに学習を促進する効果」のことである (Roediger & Karpicke, 2006b)。この直接的な学習促進の主なものとして「テスト効果」がある (e.g., Butler & Roediger, 2007; Carpenter, 2009, 2011; Carpenter & DeLosh, 2005; Carpenter & Kelly, 2012; Coppens, Verkoeijen, & Rikers, 2011; 遠藤, 2007; Marsh, Fazio, & Goswick, 2012; Roediger & Karpicke, 2006a, 2006b; Roediger, Putnam, & Smith, 2011; 多鹿, 2008)。

Roediger & Karpicke (2006a) では、記銘すべき内容を学習させる「学習フェーズ」と、最終的な記憶成績を測定する「評価テスト」からなる実験が行われている。参加者は学習フェーズにおいて、以下に示す「学習」もしくは「テスト (評価テストに先行する『事前テスト』)」によって 256—275 語の文章を学習した。「学習」では、参加者は 5 分間でできるだけ繰り返して文章を熟読するよう求められ、「テスト」では文章の内容を 10 分間で思い出せるだけ自由再生するよう求められた。学習またはテストは学習フェーズの間に計 4 回実施され、参加者は学習のみを 4 回繰り返す「SSSS 群」、学習を 3 回、その後テストを 1 回行う「SSST 群」、学習を 1 回行った後にテストを 3 回繰り返す「STTT 群」のいずれかの群に割り当てられた。つまり SSSS 群は 5 分間の文章熟読を 4 回行い、SSST 群は 5 分間の文章熟読を 3 回、内容を思い出すテストを 1 回行った。また STTT 群は 5 分間の文章熟読を 1 回行った後テストを 3 回

行った。どの群も 1 回 (5 分) の学習の間に平均して 3.5 回文章を読むことができたため、SSSS 群は 14.2 回、SSST 群は 10.3 回、STTT 群は 3.4 回文章を読むことができた。さらにテスト段階では文章全体のおよそ 70% しか再生できず、その後には正しい内容のフィードバックもなされなかったため、正しい内容の経験量は SSSS 群が最も多く、続いて SSST 群であり、STTT 群は最も少なかった。

以上の学習フェーズの 5 分後、または 1 週間後に評価テスト (再生テスト) を行った結果、5 分後のテストでは SSSS 群の正再生率が 83%、SSST 群が 78%、STTT 群が 70% と、学習の回数が多いほど成績が良かった。しかし 1 週間後の評価テストにおいては成績が逆転し、SSSS 群の正再生率が 40%、SSST 群が 56%、STTT 群が 61% であった。つまり、評価テストの前に繰り返しテストを受けることによって、情報の長期的な保持が直接促進された。

このように、事前に学習させた「記銘すべき内容」について、読んで学習する「再学習」に比べて再生や再認などを求める「事前テスト」を実施してその内容を検索させたものの方が、最終的な想起テストにおいてより多く想起される。これが「テスト効果」である。Roediger & Karpicke (2006a) は文章を刺激として使用したものであるが、単語リスト (e.g., Carpenter, 2009, 2011) や、顔と名前 (Carpenter & DeLosh, 2005) もしくは記号と単語 (Coppens et al., 2011) の対連合学習、空間的な位置の学習 (Carpenter & Kelly, 2012) 等様々な種類の刺激で生じ、また実際の教室場面でも繰り返し再現されている (e.g., Butler & Roediger, 2007; Marsh et al., 2012)。

2.2 間接効果

一方、「テストが個人の何らかのプロセスに影響し、そのプロセスを介して学習が促進される効果」はテストによる学習促進の間接効果と呼ばれている (Roediger & Karpicke, 2006b)。この間接効果については古くから言及されており、例えば Spedding, Ellis, & Heath (1870-1876 服部訳 2005) によれば、Bacon も 1620 年に発行された著書「ノヴム・オルガスム」の中で、覚えるべき文章全体を 20 回繰り返して読むよりも、暗唱しようとしながら、間違っただ箇所だけ読み返して修正することを 10 回行う方が良く覚えられると述べられている。また学校教育の場においても、前述したような様々なテストの結果から正答を学ぶよう教授者から指示されることもある。このような「事前にテストを受けると、その後の再学習がより効果的なものになる」現象や「テストによって、学習者は自分が何を知っていて何を知らないのかがわかる」、「頻繁にテストを行うことによって、学習者はより勉強するようになる」ことなどはテストによる間接的な学習促進の一つである (Roediger et al., 2011)。

事前にテストを受けることがその後の再学習 (フィードバックされた情報) の学習を促進する現象について体系的に検討したのは Izawa (e.g., 1969, 1970) である。これらの研究では実験者によって手がかりとそれに対するターゲットが対提示される「再学習」と、参加者が手がかりからターゲットを解答する(正しい反応がフィードバックされない)「テスト」が、学習方略としてどのように異なるか検討している。例えば Izawa (1970) では、手がかりとターゲット (3 文字の無意味つづりと 2 桁の数字もしくは名詞のペア) からなる刺激の学習を行わせている。分散された n 回の再学習試行の間にテスト試行を m 回実施し (例えば $n = 5, m = 5$ の場合 $R_1T_1T_2T_3T_4T_5R_2T_1T_2\cdots T_3T_4T_5R_5$, 参加者は R_1 で初めて刺激

を学習する), そのテスト試行の回数が学習に及ぼす影響について検討している。連続したテスト中の成績の変化や, 再学習の直後のテストでの成績について検討した結果, 連続したテスト中の成績は一定であったが, 再学習後のテスト成績は学習の間にテストを1度だけ行った場合より, 複数回行った場合の方が向上した(ただし繰り返すテストの回数が多ければ多だけ成績が向上するのではなく, 7回程度が最も効果的であると示唆されている)。つまり, テストはその後の再学習の効果を向上させることが明らかとなった。

3. テストにおけるエラーの効果

前述の通り, テスト効果は事前テストで検索された情報に対する直接的な学習促進効果であり, 事前テストを行ったことが後からフィードバックされた情報の学習を促進する間接的な効果とは異なる。そのため, テスト効果は基本的に「事前テストで正答したものに対する効果」であると言える。テスト効果の実験においては事前テストの後に正答をフィードバックしている研究も存在するが(e.g., Brewer & Unsworth, 2012; Kang, McDermott, & Roediger, 2007), このような手続きはテストでの検索による直接的な学習促進とフィードバックが提示されたことによる間接的な学習が混在してしまうため, できるだけ事前テストでの正答率を高くするべきだという指摘もある(Karpicke, 2017)。

しかしテストにおいては誤ることもあり, その場合は後から提示されるフィードバックの内容を正しく記憶する必要がある。それでは事前テストにおいて誤答した場合, 記憶に対してどのような影響が生じるのか。学習において誤りは避けることのできないイベントであり, 特に学習の苦手な者ほどその頻度は高くなりがちであると考えられる。テストで誤

る、つまり記銘すべき内容とは異なる内容を検索してしまうことの影響について検討することは、テストによる学習促進効果を学校教育の場に応用する際に重要な点であるといえる。

3.1 行動主義的学習理論に基づく予測と errorless 学習

テストにおける誤答が学習を促進するか否かに関する一つの考え方として、事前テストにおける誤答は後から提示される正答（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するというものがある。そのように考えられてきた主な理由として、行動主義的な学習理論がある。行動主義的学習理論では、学習段階での誤りはそれを訂正したり正答を確認したりした場合であっても「誤りの痕跡」が残るため学習を阻害すると考えられる。事前に経験した誤りに付加した「誤りであった」という情報が忘れ去られて、正答を思い出すべき時に誤って思い出されてしまう可能性がある（Skinner, 1968 村井・沼野監訳 1969）。この理論をもとに考案された学習方法が errorless 学習である。Errorless 学習はその名の通り、誤りを最小化するよう学習をいくつかの段階に細分化して行う方法である。例えば顔と名前の対連合学習の場合（三村, 1998）、誤りが生じやすいような手続き（errorless 学習に対して「errorful 学習」）では、まず顔とそれに対する名前を学習したのち、顔だけを見て名前を想起させる。一方の errorless 学習の手続きでは、まず顔とそれに対する完全な名前（例えば「たかはし」）を対提示し、その後1回目の学習では1文字（「たかは __」）、2回目は2文字（「たか __ __」）というように順に空欄にして、その空欄を想起させる。この学習方法は、現在でも記憶障害の患者に対するリハビリテーションや高齢者の学習において効果的であるといわれている（Baddeley & Willson, 1994; 三村,

1998; レビューとして Clare & Jones, 2008)。ただし近年, errorless 学習と errorful 学習の効果の違いは刺激の特性によるものであると指摘する研究もある (Bridger & Mecklinger, 2014; Cyr & Anderson, 2015)。これらの研究では, 学習内容が手がかり単語とターゲット単語のような概念学習の場合は errorful 学習が, リスト化された単語の学習のような場合は errorless 学習が効果的であると示している。

3.2 誤答の効果を検討することの問題

テストにおける誤答の効果について検証するための方法として, 例えば事前に学習した内容に対して事前テストで正答した刺激と誤答した刺激, もしくは事前テストを受けていない (再学習を行った) 刺激の評価テスト成績を比較するものが考えられる。しかしこれらの比較は「item selection effect」と呼ばれる現象によって影響を受けることから, 適当でないとされている (Pashler, Zarow, & Triplet, 2003)。事前テストで思い出せた項目は, 思い出せなかった項目と比較して, 多くの場合「参加者にとって簡単な項目であった」可能性が否定できない。一方事前テストで思い出せなかった項目は「参加者にとって難しい項目であった」可能性がある。つまり事前テストで正答した項目や誤答であった項目の評価テストにおける成績には, 事前テストを行ったことによる効果に参加者にとっての「項目の容易さ」が交絡する (Kornell, Hays, & Bjork, 2009; Kornell, & Vaughn, 2016; Pashler et al., 2003)。

前述の Izawa (e.g., 1970) の結果も, この item selection effect による影響が考えられる。つまりテストはその後の再学習の効果を向上させ得るが, これらの刺激はそもそも参加者にとって難しかった (再学習の効果が過小評価されている) 可能性がある。

3.3 誤検索効果

この *item selection effect* を受け、Kornell et al. (2009) は参加者が高い確率で誤るような事前テスト (Figure 1) を実施し、テストでの誤答がその後の情報保持に及ぼす影響を検討している。

Kornell et al. (2009) の実験 3 では、学習フェーズで手がかり単語とターゲット単語のペア (例えば「tide-beach」) の対連合学習を行わせている。この単語対は、手がかり単語からターゲット単語が連想される頻度 (連想強度) が 5% 程度のものが採用された。60 組の単語対のうち半数をテスト条件に割り当て、これらの刺激は単語対を学習させる前に手がかり単語からターゲット単語を解答させる「事前テスト」を行い、その後正しいターゲット単語をフィードバックとして提示した (Figure 1 上段)。参加者はテストの前に正しいターゲット単語を知らされておらず、また連想強度が低いため、正しいターゲット単語は手がかり単語から連想されにくい。そのため、事前テストにおいて参加者は 95% 程度の高い確率で「wave」など正しいターゲット単語とは異なる語を答えた (手がかり単語とターゲット単語の間には弱いながら関連があるため、事前テストの段階で手がかり単語「tide」からターゲット単語「beach」を正しく解答する参加者もいた。しかし事前テストの段階でターゲット単語が正しく解答された刺激は *item selection effect* の問題から後の分析からは除外された)。一方の統制条件では事前テストは行わず、手がかり単語とターゲット単語を対提示した (Figure 1 下段)。

テスト条件と統制条件の評価テスト成績を比較したところ、テスト条件の成績の方が統制条件より有意に高かった。つまり、事前テストで正答とは異なる情報を解答した場合でも学習促進が生じることが明らかとなり、この現象は誤検索 (*failed retrieval*) 効果と呼ばれている。なお、

上述の統制条件は事前テストを経験していないため、手がかり単語の提示時間がテスト条件に比べて短い。しかし同じ Kornell et al. (2009) の実験 4 では、テスト条件と統制条件の刺激提示時間を等しくしている(つまり統制条件では、テスト条件の事前テストとフィードバックの提示時間と同じ時間、手がかり単語とターゲット単語を対提示した)が、テスト条件の再生成績は統制条件に比べて高いままであった。つまり評価テストの成績差は提示時間の違いによってもたらされているのではない。

このように、テストにおいて誤答であった場合も、記銘すべき刺激を提示されるだけの条件に比べて学習が促進される。誤検索効果も刺激の種類に関わらず生じる頑健な現象であり、Kornell et al. (2009) 同様単語対を使用した研究 (e.g., Grimaldi & Karpicke, 2012; Kornell, 2014 の実験 1) 以外にも、散文を読みその内容についてのテストを行った研究 (Kornell, 2014 の実験 2, 実験 3; Richland, Kornell, & Kao, 2009) などが存在する。

第 2 節 誤検索効果のメカニズム

1. メカニズムを検討する意義

第 1 節で述べた通り、学習者は教育の場において日々テストを経験している。そのテストは、多くの場合知識や能力の評価のために実施されることが多いが、テストを受けることが知識や能力の定着を促進させることが実験によって繰り返し示されている。また、それはテストで正答した場合だけに限らず、誤答であった場合でも、その後フィードバックされた内容の学習が促進される。テストによる学習促進は、テスト自体が学校教育の場で既に使われているため、学習方略として導入するのに

コストがかからない (Roediger & Pyc, 2012) が、テストという活動を学習者にとってより良いものにするためには、テストにおけるどのような誤答が学習を促進するのかを明らかにする必要がある。前節で述べた通り、学習において誤りを完全に排除することは難しく、また学習成績の低い者ほど誤りの頻度が高いことから、誤検索効果のメカニズムを明らかにすることはより良い学習の機会をもたらすことにつながると考えられる。

では、テストにおける誤りがどのようにして学習を促進するのか。まず、テストによる学習促進の際の認知過程を概説するものとして提案されている「二段階モデル」について述べ、その後誤検索効果のメカニズムに言及する理論として「意味ネットワークの活性化」に基づいた説および「学習時の文脈」に基づいた説について述べる。

2. 二段階モデル

Kornell, Klein, & Rawson (2015) は、テストによる学習促進が生じるメカニズムを説明するために「二段階モデル」を提案している (Figure 2)。これによると、学習者はまず第一段階において、提示された手がかりから正答を検索しようとする。そして正しく正答を検索する (Figure 2 上段) か正答フィードバックが与えられる (Figure 2 下段) ことによって、第二段階で正答に対する処理が行われる。このモデルに照らせば、テスト効果も誤検索効果も以下のように説明できる。つまり、第一段階で検索しようとするのが、第二段階で検索の成功、もしくはフィードバックによって利用可能になった記録すべき解答の符号化を強化する。

このことから Kornell & Vaughn (2016) は、テスト効果と誤検索効果は表面的には異なっている点もあるが、深いレベルにおいては、つまり

第一段階の処理が第二段階の処理を促進するという点は同一であり、同じメカニズムに基づくものである可能性を指摘している。しかし一方でこの説明は検索による学習促進の際に生じている認知過程を二段階に分割しただけで、なぜ検索（第一段階）がその後の学習（第二段階）を促進するのか、そのメカニズムを実質的には説明していないとも述べている。

3. 意味ネットワークの活性化に基づく説

誤検索効果のメカニズムを具体的に説明するものとして Grimaldi & Karpicke (2012) は意味ネットワークの活性化を挙げている。この研究の実験 1 では Kornell et al. (2009) 同様の単語対を用いた手続き (Figure 1 参照) において、単語間の意味的な関連性の有無が学習促進に影響するか否かを検討した。実験では、参加者に学習させた 60 組の単語対のうち、半数は「tide-beach」のような意味的に弱い関連性のあるものを、残り半数は「pillow-leaf」のような意味的にまったく関連のないものを用いた (いずれも左側が手がかり単語、右側がターゲット単語であった)。その結果、意味的に関連のある単語対を用いた場合には従来通りの誤検索効果が生じ、事前テストを行った方が評価テスト成績が高かった。しかし意味的に関連のない単語対の評価テストでの想起成績は、事前テストを行っても向上しなかった。

さらに実験 2 では、実験 1 同様の (つまり Kornell et al., 2009 と同じ) 事前テストを行う群と事前テストを行わない統制群に加え、事前テストで誤って解答されやすいルアー単語 (手がかり単語「tide」に対する「wave」) を実験者が与えた群、ルアー語の語幹 (手がかり単語「tide」に対して「wa_」) を与えて参加者の解答に制限を設けた群の学習成績

を比較した。その結果、実験 1 同様の事前テストを行った群は学習促進が生じたが、ルアー語を提示されたり解答に制限を設けられたりした群では学習促進が生じなかった。

これらの結果は以下のように説明されている。事前テスト時に検索を行ったことによって、手がかり単語と関連する項目が活性化する。誤検索効果では参加者はターゲット単語を知らされていないため、手がかり単語から連想される語のネットワーク (**search-set**) が活性化し、更にその **search-set** とターゲット単語が関連性を持つことによってターゲット単語の符号化が促進される。そのため、単語対の間に意味的な関連性がある場合は学習促進が生じ、意味的な関連性がない場合は、手がかり単語から連想された **search-set** とターゲット単語の間にも関連性がないため、符号化が促進されない。またルアー語を提示したり、ルアー語を解答するように制限したりした場合は学習が生じなかったことから、事前テストは評価テスト時に手がかりとなるような語を単に追加しているわけではない。さらにルアー語を提示して学習させることは **search-set** が活性化せず、解答語を制限した検索を行わせることは **search-set** の活性化も制限する。つまり検索が行われずに **search-set** が活性化しなかったり、活性化が制限されたりした場合も学習促進が生じなくなる。

Grimaldi & Karpicke (2012) はさらに、**search-set** 説は情報の検索されやすさを説明する **search of associative memory model** (以下 **SAM model**) とも一致すると示唆している。**SAM model** (Gillund & Shiffrin, 1984; Raaijmakers & Shiffrin, 1981) では、学習された情報の検索されやすさは、その情報に付加している項目の量によって決まると考えられている。つまり、誤検索効果は事前テストでの検索によって手がかりに関連する **search-set** が活性化し、活性化した **search-set** と関連するターゲットが提

示されることによって、**search-set**に含まれる項目が **SAM model** であるところの「付加情報」となり、ターゲットの想起可能性を向上させる。

以上のように **search-set** 説は誤検索効果のメカニズムを **search-set**、つまり「意味的なネットワーク」の活性化によって説明する理論であるが、これはテスト効果のメカニズムを説明するために提案された **elaborative retrieval** 説を踏襲したものである。

Carpenter (2009) は、テスト効果に意味的な情報（ターゲットに関連する **elaborative information**）が及ぼす影響について検討するため、刺激が持つ意味的な関連度の高さがテスト効果に及ぼす影響について検討している。実験では、意味的な関連度の高い単語対と関連度の低い単語対（それぞれ「**Toast-Bread**」と「**Basket-Bread**」など。いずれも左側の単語が手がかり単語、右側がターゲット単語）を使用し、はじめにすべての単語対を学習させた後、半数の単語対は事前テストとして手がかり単語のみを提示してターゲット単語を検索させた（テスト条件）。残り半数は、手がかり単語とターゲット単語を対提示した（再学習条件）。その後実施した評価テストの成績について分析した結果、テストした単語の想起成績が高く、典型的なテスト効果が生じた。また関連度の高低と事前の学習方法の交互作用において、意味的な関連度の弱い単語対においてのみ、再提示条件よりテスト条件の刺激の想起成績が高かった。関連度の高い単語対（**Toast-Bread**）は、事前テストの際に手がかり単語からターゲット単語以外の情報が活性化しにくく、一方の関連度の低い単語対（**Basket-Bread**）はターゲット単語以外の情報が活性化しやすいと考えられる。そのため、この活性化した情報がターゲット単語の符号化を促進し、関連度の低い単語対の評価テスト成績が向上したと説明している。

以上の結果から、**elaborative retrieval** 説に基づくと、テスト効果が生じ

るメカニズムは以下のように考えられている。事前テスト時に手がかりをもとにターゲットとなる情報を検索することによって、手がかりと意味的に関連した情報が活性化され、この活性化した情報により評価テスト時に時に解答すべきターゲットが想起される可能性が向上するのである。

以上の通り、**search-set** 説は誤検索効果を、**elaborative retrieval** 説はテスト効果を説明しようとした理論であるが、**elaborative retrieval** 説によって誤検索効果を説明しようとする研究も存在する (e.g., Kornell, 2014; Kornell & Vaughn, 2016)。しかしこれらの研究では **search-set** 説と **elaborative retrieval** 説を同一視しているもの、**elaborative retrieval** 説は事前テストで活性化した情報が手がかりとターゲットを結ぶ媒介として機能する (cf. Carpenter, 2011) 点が **search-set** 説と異なると指摘するものもあるが、いずれの説も検索時に生じた「意味的な情報の活性化」によって、記銘すべき情報の符号化促進を説明しようとしており、意味的な情報の活性化量が多くなると学習がより促進されると想定する点で類似したものであるといえる (Kornell & Vaughn, 2016)。

4. 学習時の文脈に基づく説

Search-set 説や **elaborative retrieval** 説では、テスト効果や誤検索効果のメカニズムとして、検索による意味的なネットワークの活性化を想定している。しかし、これらのほかにもテストによる学習促進を説明する理論が存在する。二段階モデルが提案するように誤検索効果とテスト効果が同じメカニズムに基づくものである可能性を踏まえると、学習時の文脈によってテスト効果を説明する理論である「**episodic context** 説」も誤検索効果を説明しうると考えられる。

人が情報を記憶する際には、情報そのものの語彙や意味的な特徴だけではなく、学習時の一時的な文脈などの情報も記憶されると言われている (e.g., Howard & Kahana, 2001)。このことを踏まえ episodic context 説 (Karpicke, Lehman, & Aue, 2014) では、学習時と検索時で変化する文脈の符号化によって、テストによる学習の促進を説明している。例えばはじめに単語を学習して、その後事前テストを受けるテスト効果の手続きを考える。はじめの学習の際には単語だけではなく、それが学習されたときの文脈 (文脈 A) の情報も併せて符号化される。その後の事前テストでは、学習した単語を思い出すための手がかりとして文脈 A が活性化され、さらに事前テスト時の文脈 (文脈 B) も符号化されることで、思い出された単語には文脈 A と B が統合された情報が付加される。その結果、事前テスト時の検索によってアップデートされた文脈が、評価テスト時に有効な手がかりとして機能すると考えられている。しかしこの説について Kornell & Vaughn (2016) は、学習後に事前テストを実施するテスト効果は説明できても誤検索効果を説明することは難しいと述べている。

第 3 節 誤検索効果が意味ネットワークの活性化によるものである 根拠—遅延フィードバックによる影響—

意味的なネットワークの活性化によって誤検索効果のメカニズムが説明できると考えられる根拠として、以下の 2 点が挙げられる。第一に、前述のとおり、誤検索効果は意味的に関連のない単語対を刺激とした場合には生じない。これは事前テスト時の検索によって活性化したネットワークとターゲット単語の間に関連性がないため、符号化が促進されな

かったためであると考えられている。第二に遅延フィードバックの影響である。刺激として単語対を用いた場合は直後フィードバックでないと効果が生じない (e.g., Grimaldi & Karpicke, 2012 の実験 3; Hays, Kornell, & Bjork, 2013; Kornell, 2014 の実験 1) が, 文章を用いた場合は正答フィードバックまで遅延 (1 日程度) があっても学習促進が生じる (Kornell, 2014 の実験 3; Richland et al., 2009)。

例えば刺激として単語対を用いた Grimaldi & Karpicke (2012) の実験 3 の遅延フィードバック条件では, 学習フェーズを 2 つのブロックに分け, 第一ブロックですべての刺激の事前テスト (手がかり単語からターゲット単語を検索する試行) を行い, 第二ブロックで正答フィードバックを提示した。この条件の成績は統制条件と同程度であり, 学習は促進されなかった。この時の実際の遅延時間は明言されていないが, 手続きから計算すると, 第一ブロックの継続時間はおよそ 7.5 分, 第二ブロックはおよそ 5.5 分であったことから, 最長でも 13 分程度だと考えられる (なお, 実験ではそれぞれのブロック内での刺激の提示順序はランダムにされていた)。同様に単語対を用いた Hays et al. (2013) の実験 1 や Kornell (2014) の実験 1 でも遅延フィードバックによる効果の消失が再現されており, これらの実験での平均遅延時間はそれぞれ 9.5 分, 4 分程度であった。つまり, 単語対を刺激とした場合, 数分から十数分程度の短い遅延時間でも学習促進が消失する。この結果を Grimaldi & Karpicke (2012) は事前テスト時にターゲット単語を検索することによって, 短時間持続する意味的なネットワークが活性化し, それがフィードバックされた情報の学習を促進するためであると説明した。

一方, 刺激として文章 (例えば「Who was Time Magazine's "Man of the Year" in 1938?」などの問題文) を用いた Kornell (2014) の実験 3a では,

はじめに事前テストとして参加者に文章を見せ、それに解答させた。誤検索効果の手続きであるため参加者は事前に答えは知らされておらず、およそ 86%の問題に正しく解答できなかった。また実験データの分析には、これらの誤答のみが用いられた。事前テストを実施した 24 時間後に正答フィードバックを与え、それからさらに 24 時間後（事前テストの 48 時間後）に評価テストを行ったところ、遅延フィードバックにもかかわらず事前テストによる学習促進効果が生じた。この結果から Kornell (2014) は、直後フィードバックでしか学習促進が生じない単語対による誤検索効果も、遅延フィードバックでも学習促進が生じる文章を用いた誤検索効果も、学習促進は検索を行ったために生じたことに違いはないが、刺激として文章が提示されると、単語対の場合に比べてより多くの情報が活性化される。この活性化する情報量の違いが、遅延フィードバックでの成績差をもたらしたと説明している。

つまり Grimaldi & Karpicke (2012) や Hays et al. (2013), Kornell (2014) の結果を総合すれば、誤検索効果の生起には意味的なネットワークの活性化が必要であり、かつ活性化される情報の量が増えればフィードバックまでの遅延時間が長くなっても学習が促進されると考えられる。しかし単語対のみを刺激として活性化する意味ネットワークにおける情報の活性化量（SAM model にもとづけば、記録すべき情報に付加したアイテム量）を操作し、誤検索効果の生起パターン、特に遅延フィードバック時の学習促進効果について検討したものはない。

もし単語対の学習で意味的なネットワークにおいて活性化される情報の量を増やしても遅延フィードバックで学習促進が生じないとすれば、誤検索効果には意味ネットワークにおける情報の活性化量のみでは説明できない現象が含まれるといえる。また誤検索効果を学校教育の場に応

用する場合、学校では児童・生徒の誤答にすぐフィードバックが与えられるとは限らない。どのように誤答が生成されれば遅延フィードバックであっても学習促進効果が生じるのかについて明らかにすることは、学校での教育実践への応用という視点において重要であると考えられる。

第4節 本研究の目的

以上を踏まえ、本研究では遅延フィードバック条件下でも意味的なネットワークを活性化させることによって誤検索効果が生じるかを検討し、誤検索効果が生起するメカニズムを明らかにすることを目的とした。ここまで挙げた先行研究の結果を総合すれば、誤検索効果の生起には意味的なネットワークの活性化が必要であり、かつ、たとえ刺激が単語対であった場合でも、ネットワークにおいて活性化する情報量が増えればその効果は増大する（特にフィードバックまでの遅延時間が長くなっても学習が促進される）と考えられる。

検索時に活性化する情報量がテスト効果、もしくは単語対以外を使用した誤検索効果に及ぼす影響について示唆する研究はいくつか存在する。例えば Lehman & Karpicke (2016) の実験 3a-5 では、一度単語対を学習した後の事前テストで（つまりテスト効果の手続きで）、手がかり単語と意味的に関連し、かつ覚えるべきターゲット単語とは異なる語（媒介語）を複数生成させた。その結果、手がかり再生を求めた評価テストではむしろ媒介語の個数が増えると成績が低下した。また誤検索効果については、刺激として文章（トリビア問題）を使用した Vaughn, Hausman, & Kornell (2017) がある。この研究では、問題提示後から正答フィードバック提示までの検索時間を 0 秒（検索はさせず、問題と解答を対提示し

た), もしくは 5 秒—30 秒と操作した。その結果, 検索時間を長くすると事前テストの正答率は高くなった (つまり参加者は正答を考え続け, 意味的な活性化もより多く生じたと考えられる) が, 評価テストの成績には「検索をしたかどうか」しか影響せず, 時間の長さによる影響はなかった。

以上の研究からは活性化する情報量の違いは学習促進にそれほど大きな影響をあたえないようにも考えられる。しかし Lehman & Karpicke (2016) では覚えるべき単語とは異なる語を参加者に生成させており, これらの語が干渉を生じた可能性がある。また Vaughn et al. (2017) では刺激として文章が使用されており, Kornell (2014) を踏まえれば短時間でも十分豊かな情報が活性化したとも考えられる。

以上を踏まえ, 研究 1 における仮説は以下の 3 つである。

1. これまでの先行研究同様, テストで誤ることが学習を促進するならば, 正しいターゲット単語が提示される前に事前テストで手がかり単語からターゲット単語になりそうな語 (以下「検索語」) を検索し解答した検索条件の単語対は, 事前テストを行わない統制条件の単語対より評価テストでの再生成績が高くなる (典型的な誤検索効果が生じる)。
2. この誤検索効果は直後フィードバックの場合には生じるが, 特に検索語を一つだけ解答させる場合は, 遅延フィードバック条件下では生じない。
3. 誤検索効果が意味ネットワークの活性化によるものであれば, 検索語 (誤情報) を複数回答させる条件は, 一つだけ解答させる条件 (先行研究と同様の手続き) に比べて評価テスト成績が向上する。ただし直後に正答をフィードバックする場合, 一つだけ解答させ

る条件でも十分な誤検索効果が生じることが予想されるため、この影響は遅延フィードバック条件においてより顕著になる。一方複数解答条件でも遅延フィードバック条件下で学習促進が生じなければ、それは「ネットワークにおいて活性化する情報量が増えれば学習促進効果が増大する」という仮説が正しくない可能性を示唆するであろう。

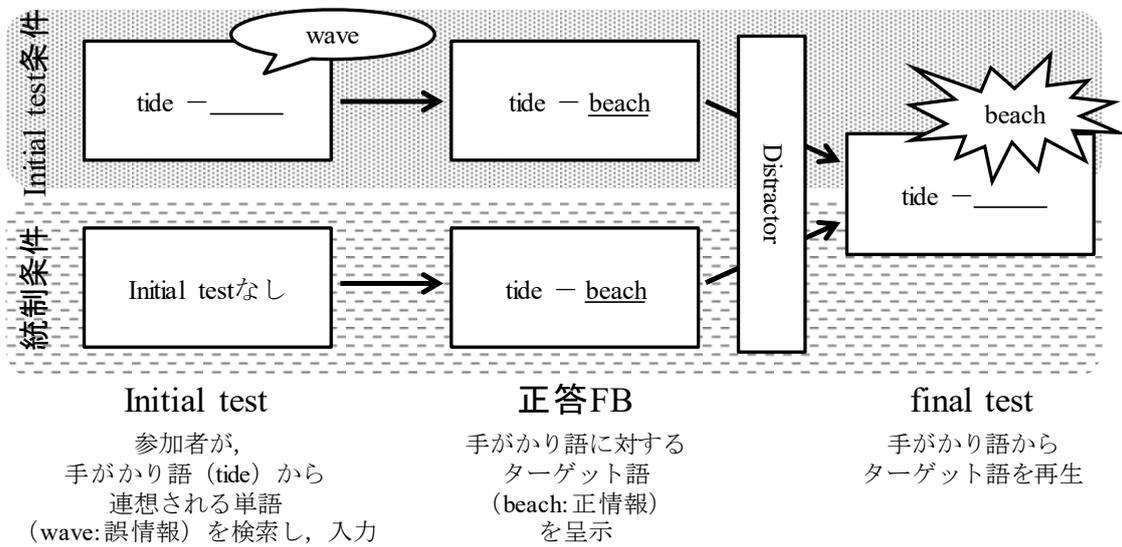
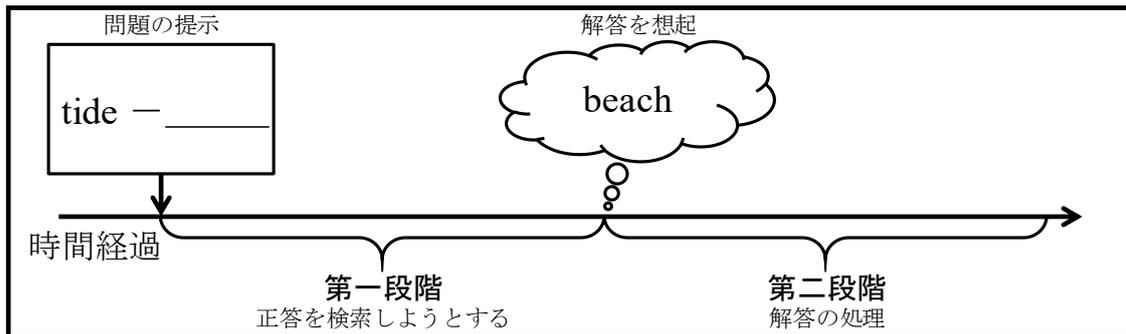


Figure 1. Kornell et al. (2009) の実験 3 の手続き

正答を検索できた場合



正答が検索できなかった場合

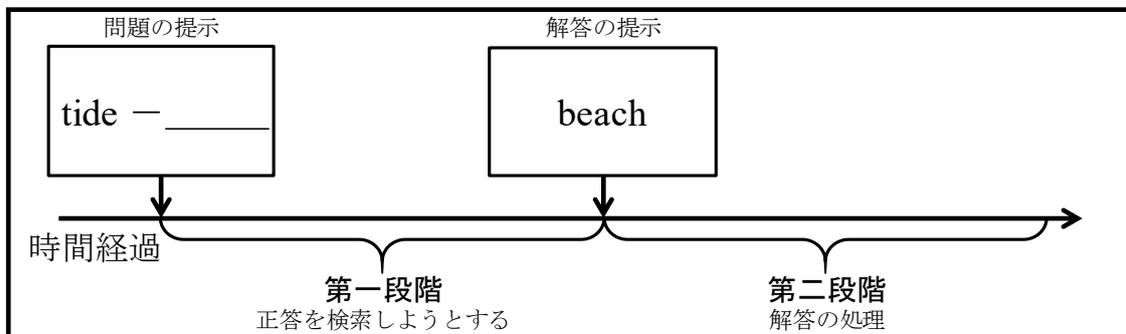


Figure 2. Karpicke et al. (2014) による、検索時の情報処理の二段階モデル。図は Kornell & Vaughn (2016) をもとに作成した。

第 2 章 誤情報の検索が正情報の学習を促進するメカニズムの検討

第 1 節 検索する情報量の影響（研究 1）

1. 研究 1-1

1.1 目的

本研究では、Grimaldi & Karpicke (2012) の実験 3 や Hays et al. (2013), Kornell (2014) の実験 1 で学習促進効果が見られなかった遅延フィードバック条件において、手がかり単語から検索する語を増やす、つまり意味ネットワークにおいて活性化する情報量を増やすことによって学習促進が生じるか検討した。

1.2 方法

参加者 大学生，短期大学生 27 名（男性 10 名，女性 17 名，平均年齢 19.9 歳， $SD = 1.2$ ）が参加した。参加者には謝金として 1000 円を支払い，参加者の扱いは APA ethical guidelines に従った。参加者の母語は日本語であり，一人ひとりの参加者は，後述する 5 つすべての条件で単語対を学習した。

刺激 手がかり単語とターゲット単語の単語対 100 組を使用した（例えば「赤字－借金」，「バイク－タイヤ」など。下線語はターゲット単語であった。すべての刺激は Appendix に記載した）。単語対は Kornell et al. (2009) を参考に，連想強度（手がかり単語が提示されたときに，1 つ目の反応としてターゲット単語を解答する割合）が .041— .054 となるよう，水野（2011）をもとに選択した。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，手がか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。

条件 単語対の学習段階（学習フェーズ）では，手がかり単語のみを

提示しての事前テストとフィードバックを行った。事前テストでの学習方法には、手がかり単語からターゲット単語となるような単語（以下「検索語」）を1つだけ解答する条件と2つ解答する条件（それぞれ単一条件、複数条件）、また事前テスト後すぐに正答（覚えるべきターゲット単語）をフィードバックする条件と、遅延後にフィードバックする条件（それぞれ直後条件、遅延条件）を組み合わせた4条件（それぞれ単一直後条件、複数直後条件、単一遅延条件、複数遅延条件）に、事前テストを行わず検索語を解答しない統制条件を加えた計5条件を設定した（Figure 3）。

単一直後条件ではまず、パソコンのディスプレイ上に凝視点としてプラスの記号を1つ、1秒間提示した。その後手がかり単語と長方形の枠を提示し、7秒以内にできるだけ早く手がかり単語からターゲット単語になるような単語（検索語）を口頭で1つだけ解答するよう求めた。単語を解答した段階で実験者がエンターキーを押して次の画面に進み、そこで正しいターゲット単語を5秒間提示した。なお7秒の間に単語を解答できなかった場合は、自動的に次の画面（本条件では正しいターゲット単語の表示）に移った。参加者には手がかり単語とターゲット単語をペアにして記憶するよう指示した。

複数直後条件では、凝視点としてプラスの記号を2つ、1秒間提示した後、単一直後条件同様手がかり単語と長方形の枠を7秒間提示してできるだけ早く検索語を口頭で1つ解答するよう求めた。参加者が検索語を解答するか7秒経過後に次の画面に進んだ。前述の単一直後条件と異なる点として、この条件では単語を2つ答えさせるためにこの手続きを2回繰り返した（つまり解答、もしくは7秒経過後に再度手がかり単語と長方形の枠を7秒間提示して、先ほどとは異なる検索語を1つ解答す

るよう指示した)。2回目の検索語解答後は、単一直後条件同様に手がかり単語と正しいターゲット単語を5秒間提示し、それらをペアにして記憶させた。

単一遅延条件は、事前テストは単一直後条件と同様であった。異なっただ点として、参加者が単語を1つ解答するか7秒経過後、すぐには正しいターゲット単語のフィードバックを行わず、次の刺激の条件に対応した凝視点を提示した。フィードバックは一定数の単語の学習が終了した後に、後述する統制条件と同じ方法で提示した。

複数遅延条件も、事前テストは複数直後条件と同様であり、フィードバックは単一遅延条件と同様、一定数の単語の学習が終了した後に統制条件と同じ方法で提示した。

統制条件は、凝視点としてアスタリスクの記号1つを1秒間提示し、事前テストは実施せずに手がかり単語とターゲット単語だけを5秒間提示した。

これらの条件へはそれぞれ20単語ずつを割り当て、割り当て方はカウンターバランスをとった。

手続き Kornell et al. (2009) 同様、実験は学習フェーズと評価テストで構成した。Kornell et al. (2009) と異なる点として、フィードバックに遅延を設定するため、Kornell (2014) を参考に学習フェーズをプログラム上で4つのブロックに分割し、遅延条件の事前テストとフィードバックを隣り合う2つのブロックに分けて実施した (Figure 4)。つまり、第 n ブロックで手がかり単語が提示された単一遅延条件、および複数遅延条件の単語対は、続く第 $n + 1$ ブロックで正しいターゲット単語のフィードバックを行った。ブロックの間に休憩は行わず、参加者にも学習フェーズがブロックに分かれていることは説明しなかった。遅延条件にお

いて、手がかり単語提示からターゲット単語提示までの遅延時間は 0.43—11.65 分 ($M = 5.63$ 分, $SD = 2.39$ 分) であった。

各ブロックにおいて、それぞれの条件で学習する単語対は 5 組ずつであり、計 25 組の単語対をランダムな順序で提示した。なお第 2 ブロック以降は遅延条件のフィードバック試行（単一遅延条件 5 組、複数遅延条件 5 組）を加えた 35 組をランダムに提示した。また第 4 ブロックの遅延条件に割り当てられた単語対 10 組は直後条件と同様の手続きで提示し、分析からは除外した。加えて、第 1 ブロックで単一直後条件と複数直後条件、統制条件で学習した 5 組ずつ 15 組の単語対も、フィードバックから評価テストまでの経過時間を遅延条件と等しくするため分析から除外した。

参加者から実験参加の同意を書面で得た後、学習フェーズでの各条件の学習方法について教示した。その後、評価テストでは使用しない 20 組の単語対を用いて練習を行い、教示を理解しているかどうか確認した。確認後、学習フェーズを行った。学習フェーズでは、前述の単一直後条件、複数直後条件、単一遅延条件、複数遅延条件、統制条件の 5 つの方法で単語対を学習させた。ブロック内での単語対の提示順序は参加者ごとにランダムであった。

学習フェーズ後、評価テストの前にディストラクタを実施した。ディストラクタでは、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を提示し、参加者にはテンキーを用いて解答を画面上の枠の中に入力するよう指示した。問題（数式）は 200 個のリストの中からランダムに提示した。また問題提示後 1 分以内に解答が入力されなかった場合は、自動的に新たな問題を提示した。

評価テストでは手がかり再生課題を行った。ディスプレイ上に 1 秒間

凝視点（“+”）を提示した後，手がかり単語と枠を 7 秒間提示し，参加者にはその間に手がかり単語と対になっていた正しいターゲット単語を思い出し，口頭で解答するよう指示した。解答時間は 7 秒間で固定し，ターゲット単語が思い出せなかったり誤って解答したりした場合も正答は提示しなかった。なお，評価テストの手がかり再生課題は学習フェーズ時の学習条件によらず同一の手続きであり，手がかり単語の提示順序は参加者ごとにランダムであった。

1.3 結果と考察

第 1 ブロックにおいて単一直後条件，複数直後条件，統制条件で学習した単語対，および第 4 ブロックにおいて単一遅延条件，複数遅延条件に割り当てられていた（直後条件と同じ方法で学習した）単語対計 25 組と，先行研究（e.g., Kornell et al., 2009）同様事前テスト時に正答した単語対（4%）を除いた試行を分析対象とした。

検索を行った 4 条件と統制条件の評価テスト成績を比較した結果，検索を行った単語対の方が正当率は高く，先行研究と一致して，本研究においても事前に誤情報を検索することが正情報の学習を促進した（ $t(26) = 4.11, p < .01, d = 0.73$ ）。¹

また，4 つの検索条件について 2（遅延の有無）×2（解答数）の二要因分散分析を行った結果，遅延の有無の主効果（ $F(1, 26) = 0.39, p = .54, \eta^2_p < .02$ ）も，解答数の主効果（ $F(1, 26) = 0.09, p = .77, \eta^2_p = .01$ ）も，遅延の有無と解答数の交互作用（ $F(1, 26) = 0.04, p = .83, \eta^2_p < .01$ ）も

¹ 分散の正規性を満たしていない変数が含まれていたため，逆正弦変換を行ったものに対して同様の分析を行った結果，分析結果の傾向は変わらなかった（ $t(26) = 4.09, p < .01, d = 0.70$ ）。

有意ではなかった (Table 1)。²

以上のことから、先行研究通り事前テストで検索を行うことによって正しい情報の学習が促進された。ただし、遅延フィードバックの影響については従来の結果 (e.g., Grimaldi & Karpicke, 2012; Hays et al., 2013; Kornell, 2014) とは異なり、正答フィードバックまでに遅延があっても学習が生じることが示された。Grimaldi & Karpicke (2012) は遅延フィードバックの条件下で誤検索効果による学習促進が生じないことから「短時間のみ活性化する意味ネットワーク」が誤検索効果のメカニズムとして存在していることを示唆しているが、本実験の結果はそれを支持しなかった。また解答させる単語数を増やして意味ネットワークにおいて活性化する情報量を増やしても、誤検索効果への影響はなかった。このことから、文章を刺激として用いた誤検索効果において、遅延フィードバックでも学習促進が生じるのは、必ずしも単語対に比べてより豊かなネットワークが生じるため (Kornell, 2014) とは言えないと考えられる。

しかし本実験では「遅延条件のフィードバックを、事前テストの次のブロックで提示する」という手続きを用いたため、フィードバック提示までの時間間隔が 1 分弱から 12 分弱と大きなばらつきがあった。本実験の手続きは、遅延フィードバックによって誤検索効果が消失した先行研究 (Kornell, 2014 の実験 1) を踏襲したものであり、遅延時間も先行研究と大きな差はなかったが、このフィードバックまでの時間間隔のばらつきが剰余変数として影響した可能性は否定できない。そのため、続

² 逆正弦変換を行ったものに対して同様の分析を行った結果、分析結果の傾向は変わらなかった (遅延の有無の主効果 $F(1, 26) = 0.60, p = .45, \eta^2_p = .02$; 解答数の主効果 $F(1, 26) < 0.01, p = .98, \eta^2_p < .01$; 遅延の有無と解答数の交互作用 ($F(1, 26) = 0.30, p = .59, \eta^2_p = .01$))。

く研究 1-2 では条件内での遅延時間のばらつきを小さくし，フィードバック提示までの間隔の長短が学習促進に及ぼす影響について検討した。

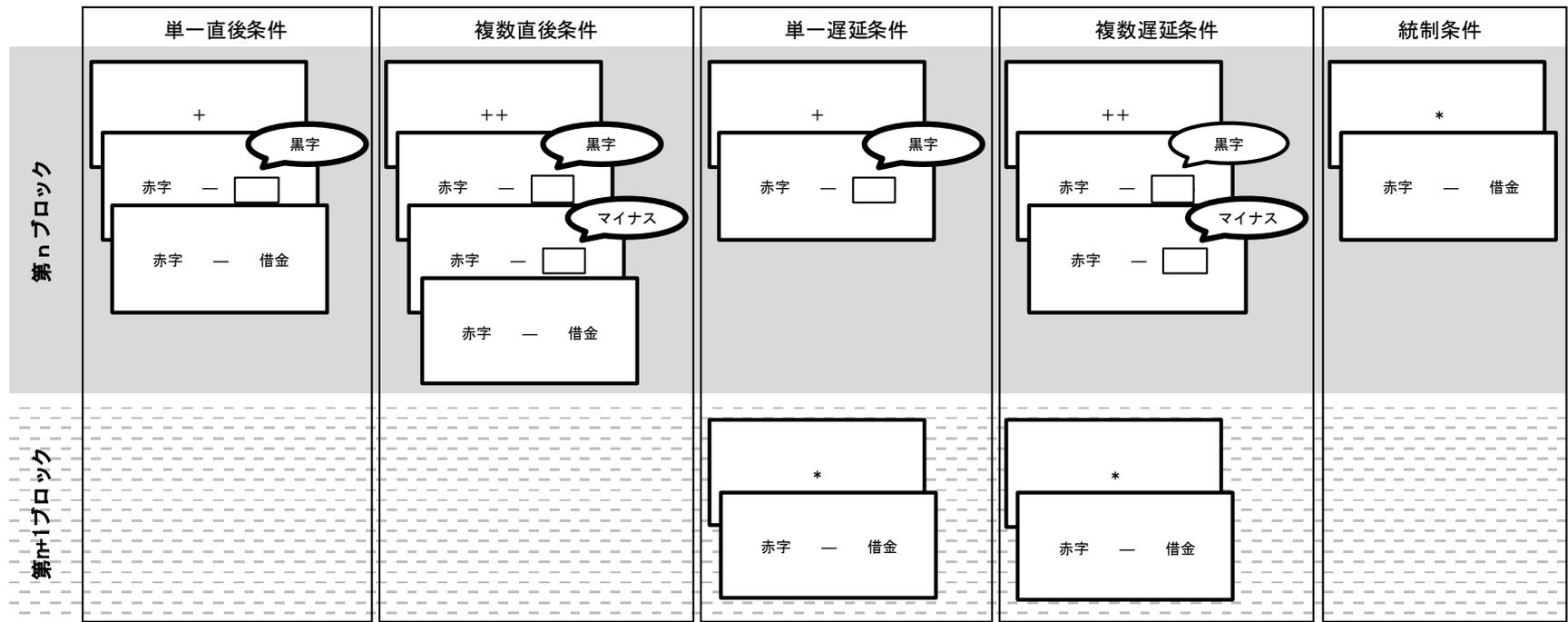


Figure 3. 研究 1-1 における各条件の刺激提示順序

	事前テスト	直後FB	遅延FB
単一直後条件	7秒×1	5秒	
複数直後条件	7秒×2	5秒	
単一遅延条件	7秒×1		5秒
複数遅延条件	7秒×2		5秒
統制条件		5秒	
	第nブロック		第n+1ブロック

Figure 4. 各条件の、ブロックごとの事前テストとフィードバックの実施順序。図中の「FB」はフィードバックを示す。

Table 1

研究 1-1 の各学習条件の評価テスト成績

	平均	<i>SD</i>
統制条件	.62	.21
単一遅延条件	.74	.21
複数遅延条件	.74	.15
単一直後条件	.75	.18
複数直後条件	.76	.16
検索条件全体	.75	.14

2. 研究 1-2

2.1 目的

研究 1-1 では、先行研究と異なりフィードバックまでに遅延があっても学習促進が生じたが、フィードバックの提示間隔が 1 分弱から 12 分弱と大きなばらつきがあった。そこで研究 1-2 では遅延条件のフィードバック提示間隔を検索後 4—5 分(短遅延条件), 10—11 分(長遅延条件)に固定し、フィードバックの遅延が学習促進に及ぼす影響について検討した。

2.2 方法

参加者 日本語を母語とする大学生, 短期大学生合わせて 25 名(男性 5 名, 女性 20 名, 平均年齢 19.32 歳, $SD = 1.97$) が実験に参加した。参加者には謝金として 1000 円を支払った。一人ひとりの参加者は、後述する 4 つすべての条件で単語対を学習した。

刺激 研究 1-1 と同じ手がかり単語とターゲット単語の単語対 100 組を使用した。

条件 参加者が誤情報を検索した後 0 分(直後), 4—5 分, 10—11 分後に正答フィードバックを行う 3 つの検索条件(直後条件, 短遅延条件, 長遅延条件)と, 検索を行わない統制条件の計 4 つの学習条件を設けた。

直後条件は研究 1-1 の単一直後条件と完全に同一であった。短遅延条件と長遅延条件は, 研究 1-1 の単一遅延条件とほぼ同一であったが, 遅延時間の設定方法のみが異なった。短遅延条件の場合は事前テスト実施から 4—5 分の間で, 長遅延条件は 10—11 分の間で, 刺激ごとにランダムに設定した時間が経過した後にフィードバックを実施した。

長遅延条件の遅延時間を確保するため, 統制条件と直後条件に 31 組,

短遅延条件に 20 組，長遅延条件に 18 組ずつ割り当てた。各条件ともそのうち 18 組を分析対象とし，残りは遅延時間を確保するためのフィラー一試行とした。刺激の条件への割り当てはカウンターバランスをとり，フィラー一試行はランダムに決定した。

手続き 研究 1-1 同様，実験は事前テストとフィードバックを行う学習フェーズと評価テストで構成し，本実験では学習フェーズをプログラム上で 6 つのブロックに分割した。

研究 1-2 の刺激提示フローを Figure 5 に示した。長遅延条件の事前テストは第 1—第 3 ブロックで提示し，短遅延条件の事前テストは第 2—第 4 ブロックで提示した。これらの条件のフィードバックは，それぞれ事前テストの 10—11 分後，もしくは 4—5 分後に提示した。加えて，直後条件（事前テストとフィードバック）および統制条件の刺激は第 1—第 6 ブロックで提示した。上述の通り，各条件のうち 18 刺激（合計 72 刺激）を分析対象とし，残りの計 28 刺激（統制条件と直後条件のうち各 13 刺激，短遅延条件の 2 刺激）はフィラー一試行とした。

研究 1-1 同様，参加者から書面で同意を得た後，学習フェーズの練習を行ってから単語対を学習させ，その後 5 分間のディストラクタを行った後に評価テストを実施した。

2.3 結果と考察

研究 1-1 同様，上述のフィラー一試行 28 組と，検索語とターゲット単語が同一であった試行（2%）を除いたものを分析対象とした。

検索を行った 3 条件と統制条件の評価テスト成績を比較した結果，検索を行った単語対の方が正当率は高く，本研究においても事前に誤情報を検索することが正情報の学習を促進した ($t(24) = 7.38, p < .01, d = 1.48$)。

³また、単語対を4つの学習条件に分けて評価テストの成績について分散分析を行ったところ、主効果が有意であった ($F(3, 72) = 17.03, p < .01, \eta^2_p = .42$)。多重比較を行った結果、3つの検索条件のそれぞれと統制条件の差が有意であった ($ps < .01$)。検索条件間に有意差はなかったが、正答率は直後条件が最も高く、続いて短遅延条件、長遅延条件の順であった (Table 2)。⁴

以上の結果より、本実験においても先行研究 (e.g., Grimaldi & Karpicke, 2012; Hays et al. 2013; Kornell, 2014) とは異なり正答フィードバックまでに遅延があっても学習は促進され、それは遅延時間のばらつきを統制しても変わらず生じた。ただし、統計的には有意でないものの、遅延時間が長くなると評価テストの成績が低下する傾向がみられた。つまり、より遅延時間を延長すると、評価テスト成績が低下する可能性が示された。

³ 分散の正規性を満たしていない変数が含まれていたため、逆正弦変換を行ったものに対して同様の分析を行った結果、分析結果の傾向は変わらなかった ($t(24) = 7.42, p < .01, d = 0.97$)。

⁴ 逆正弦変換を行ったものに対して同様の分析を行った結果、分析結果の傾向は変わらなかった (主効果 $F(3, 72) = 13.75, p < .01, \eta^2_p = .36$; 多重比較の結果、3つの検索条件のそれぞれと統制条件の差は $ps \leq .01$ で有意であった)。

ブロック	1	2	3	4	5	6
長遅延条件	IT (6 刺激)			FB		
		IT (6 刺激)		FB		
			IT (6 刺激)		FB	
短遅延条件		IT (2 刺激)		FB		
			IT (9 刺激)		FB	
				IT (9 刺激)		FB
直後条件	IT, FB (2 刺激)	IT, FB (2 刺激)			IT, FB (2 刺激)	IT, FB (7 刺激)
			IT, FB (6 刺激)	IT, FB (6 刺激)	IT, FB (6 刺激)	
統制条件	統制 (2 刺激)	統制 (2 刺激)			統制 (2 刺激)	統制 (7 刺激)
			統制 (6 刺激)	統制 (6 刺激)	統制 (6 刺激)	

Figure 5. 研究 1-2 の各条件の刺激提示フロー

Note. 各条件の刺激を，第 1—第 6 ブロックに配置した。長遅延および短遅延条件のフィードバックは，それぞれ 10—11 分，4—5 分後に提示した。1) 長遅延，短遅延の各条件のフィードバック提示までの遅延時間を確保すること，2) 参加者が，実験プログラムがブロックに分かれていることに気づかないようにするための以上 2 つの理由から，フィラー試行（図中グレーのセル）を設定した。表中の「IT」は事前テストを，「FB」はフィードバックを示す。

Table 2

研究 1-2 の各学習条件の評価テスト成績

	平均	<i>SD</i>
統制条件	.56	.16
直後条件	.73	.16
短遅延条件	.70	.17
長遅延条件	.68	.19
検索条件全体	.70	.16

3. 研究 1-3

3.1 目的

研究 1-2 より、フィードバックまでに 10 分程度の遅延があっても、また遅延時間のばらつきを統制しても、遅延フィードバック条件下でも誤検索効果が見られた。しかし遅延時間が長くなるほど、評価テスト成績が下がる傾向がみられた。そこで研究 1-3 ではフィードバック提示までの時間を事前テストの 15—25 分後に延長し、遅延フィードバックが誤検索効果へ及ぼす影響について再度検討した。

3.2 方法

参加者 母語を日本語とする大学生，短期大学生 25 名（男性 3 名，女性 22 名，平均年齢 19.6 歳， $SD = 1.4$ ）が実験に参加した。彼らには実験参加に対する謝金として 1000 円を支払い，一人ひとりの参加者は，後述する 5 つすべての条件で単語対を学習した。

刺激 研究 1-1，1-2 と同一の 100 組の単語対を使用した。

条件 研究 1-1 同様，検索を行う 4 条件（単一直後条件，複数直後条件，単一遅延条件，複数遅延条件）と，検索を行わない統制条件の計 5 条件を設定した。刺激は各条件に 20 組ずつ割り当て，割り当て方はカウンターバランスをとった。

手続き 実験はおおよそ研究 1-1 と同一であったが，異なる点として，遅延時間を設けるために Grimaldi & Katpicke（2012）の手続きを参考に学習フェーズを 2 つのブロックに分割した（Figure 6）。第 1 ブロックでは遅延条件の事前テストを，第 2 ブロックでは遅延条件のフィードバックと直後条件の事前テストとフィードバック，統制条件を実施した。第 2 ブロックは，第 1 ブロック開始の 15 分後から始め，第 1 ブロック終了

から第2ブロック開始までの間は Kornell (2014) 同様にテトリスを実施した (テトリスのプレイ時間は第一ブロック実施にかかった時間によるため、参加者ごとに異なった)。実際の遅延時間は 15.1 分—25.2 分 (平均 20.2 分, $SD = 0.5$) であった。ディストラクタと評価テストの実施方法はこれまでの研究と同一であった。

3.3 結果と考察

本実験ではフィルター試行は設けなかったため、検索語とターゲット単語が同一であった試行 (5%) のみを分析から除外した。検索条件 (4 条件の成績をプールしたもの) と統制条件の評価テスト成績を比較した結果、検索を行った単語対の方が正答率は高く、本研究においても事前に誤情報を検索することが正情報の学習を促進した ($t(24) = 5.96, p < .01, d = 0.88$)。⁵

また、検索条件について 2 (遅延の有無) \times 2 (解答数) の二要因分散分析を行った結果、遅延の有無の主効果 ($F(1, 24) = 1.34, p = .26, \eta^2_p = .05$) も解答数の主効果 ($F(1, 24) < 0.01, p = .98, \eta^2_p < .01$) も交互作用 ($F(1, 24) = 0.03, p = .87, \eta^2_p < .01$) も有意ではなかった (Table 3)。

⁶これより、本研究の手続きでは先行研究とは異なり正答フィードバックまでに平均 20 分程度の遅延があっても学習促進が生じることが明らか

⁵ 分散の正規性を満たしていない変数が含まれていたため、逆正弦変換を行ったものに対して同様の分析を行った結果、分析結果の傾向は変わらなかった ($t(24) = 5.86, p < .01, d = 0.86$)。

⁶ 逆正弦変換を行ったものに対して同様の分析を行った結果、分析結果の傾向は変わらなかった (遅延の有無の主効果 $F(1, 24) = 1.54, p = .23, \eta^2_p = .06$; 解答数の主効果 $F(1, 24) = 0.02, p = .90, \eta^2_p < .01$.; 遅延の有無と解答数の交互作用 ($F(1, 24) = 0.12, p = .74, \eta^2_p = .01$)。

となった。

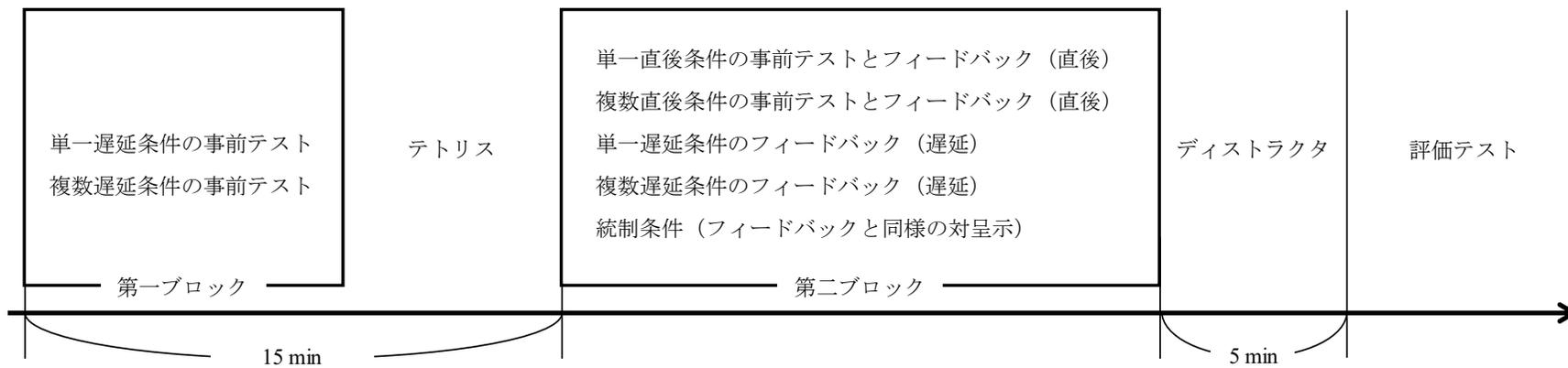


Figure 6. 研究 1-3 の全体的な流れ

Table 3

研究 1-3 の各学習条件の評価テスト成績

	平均	<i>SD</i>
統制条件	.52	.20
単一遅延条件	.67	.18
複数遅延条件	.67	.18
単一直後条件	.70	.19
複数直後条件	.70	.19
検索条件全体	.69	.16

4. 研究 1 のまとめ

これまでの結果より，仮説 1 の「事前テストを行う検索条件の単語対は，統制条件のものよりも評価テストでの再生成績が高くなる」という誤検索効果は繰り返し確認されたが，仮説 2 の「先行研究で見られた遅延フィードバック条件下での誤検索効果の消失」が，本研究では再現されなかった。また，仮説 3 の「複数解答条件は単一解答条件に比べて評価テスト成績が向上する」という結果は確認されず，また「手がかり単語から複数の検索語を解答させることによって，遅延フィードバックでも誤検索効果が生じる」という結果も確認されなかった。つまり複数解答させることによって意味的なネットワークにおける情報量を増やしても，先行研究（e.g., Grimaldi & Karpicke, 2012; Hays et al., 2013; Kornell, 2014）で示されてきた遅延フィードバックによる誤検索効果への影響はみられず，むしろフィードバックのタイミングによらず学習促進が生じ得ることが示唆された。

先行研究と本研究の手続きの異なる点の一つとして検索語の解答方法が挙げられる。先行研究では参加者に検索語をタイピングによって解答させていたが，研究 1 では口頭で解答させた。このような手続きを用いた理由は，タイピングの正確性や速度に大きな個人差があること（岩田，2015），テスト効果における解答方法に関する研究より，口頭解答とタイピング解答間に差が生じないと考えられたためであった（Putnam & Roediger, 2013）。

テスト効果とは異なるが，解答方法がその後の成績に影響するものとして産出効果がある（Forrin, MacLeod, & Ozubko, 2012）。産出効果とは，声に出さず読んだだけの学習項目より，声を出して読んだ学習項目の方が記憶されやすいという現象である（MacLeod, Gopie, Hourihan, Neary, &

Ozubko, 2010) が, Forrin et al. (2012) の実験 2 では, 記銘すべき内容を声に出すことは, 同じ内容を書いたり小声でささやいたりした場合よりも最終的な評価テストの成績を向上させることが示されている。この結果は, 声に出すことによって聴覚的な処理がなされ, 情報の符号化が促進されたために生じたと説明されている。したがって, これまでの研究 1 では事前テスト時に口頭で解答したことによって検索語の符号化が促進されたために, 遅延フィードバック条件において誤検索効果が生じなかった可能性がある。

Pinet & Nozari (2018) は, 口頭で反応する場合と異なり, タイピングによる反応時には, 反応する語を決定した後の *post-lexical* な処理が必要になることを示している。また *spelling* 時の処理について Rapp & Fischer-Baum (2014) は以下のように説明している。まず入力された情報をもとに, 長期記憶内の意味表象にアクセスする。その後, 長期記憶から引き出された正書法に関する表象がワーキングメモリシステムで処理される。ワーキングメモリではその後筆記が求められたときのために綴りに関する情報を活性化させ続け, 文字を正しい順序で書けるようにする。これを踏まえれば, 誤検索効果においてもタイピング時にはワーキングメモリの働きが必要となり, 先行研究はそちらへの注意分割が必要であったために, 本来の学習促進効果が低く見積もられていた可能性がある。

検索は注意配分の際に優先的に処理され, 並行して副課題を実施してもその影響を受けないが, 符号化は副課題下では処理が阻害される (Craig, Govoni, Naveh-Benjamin, & Anderson, 1996; Mulligan, 2008)。テスト効果においては, 事前テスト時に副課題として聴覚提示された数字の奇数・偶数判断課題を課しても成績は低下しなかったことから, 検索プロセスが重要であると考えられている (Buchin & Mulligan, 2017)。しか

し誤検索効果に対する注意分割の影響についてはほとんど検討されていない。文章刺激の場合遅延フィードバックでも学習促進効果が生じることを示した Kornell (2014) の実験 2, 3a では、事前テストでの問題と誤答が長期記憶に保存された可能性が指摘されていることから、誤検索効果、特に遅延フィードバック状況下で誤検索効果が生じるか否かは意味的なネットワークにおける情報の活性化量ではなく、検索した情報の符号化によるものである可能性もある。

以上を踏まえ、研究 2 では誤検索効果における遅延フィードバックの影響についてさらに詳細に検討するため、検索語をタイピングで解答させた場合の誤検索効果への影響について検討した。

第 2 節 検索情報の解答方法の影響（研究 2）

1. 目的

研究 1 より，検索語を口頭解答させた場合は遅延フィードバックでも誤検索効果が生じる可能性が示唆された。そこで研究 2 では研究 1-3 と同様の手続きを用い，検索語の解答方法が誤検索効果に及ぼす影響について検討した。

2. 方法

参加者 母語を日本語とする大学生，短期大学生 25 名（男性 2 名，女性 23 名，平均年齢 20.0 歳， $SD = 1.2$ ）が実験に参加した。

刺激と条件 研究 1-3 と同一であった。

手続き 手続きは，以下に示すとおり事前テスト時の解答方法をタイピングに変更したこと以外は研究 1-3 と全く同一であった。参加者には文字の変換に関わるキーを使用できないようにしたキーボードを使用させ，検索語をローマ字入力（ディスプレイ上はかな表示）でテキストボックス内にタイプさせた。タイプ後はエンターキーを押して次の試行に進むよう指示した。遅延条件の実際の遅延時間は 15.1 分—25.3 分（平均 20.0 分， $SD = 0.5$ ）であった。

3. 結果と考察

検索語とターゲット単語が同一であった試行（4%）は分析から除外した。検索条件（4 条件の成績をプールしたもの）と統制条件の評価テスト成績を比較した結果，検索を行った単語対の方が正答率は高く，本研究においても事前に誤情報を検索することが正情報の学習を促進した ($t(24) = 3.87, p < .01, d = 0.54$)。

また検索条件について 2 (遅延の有無) × 2 (解答数) の二要因分散分析を行った結果, 遅延の有無の主効果 ($F(1, 24) = 6.19, p = .02, \eta^2_p = .21$) と遅延の有無と解答数の交互作用 ($F(1, 24) = 9.24, p < .01, \eta^2_p = .28$) が有意であった。解答数の主効果は有意ではなかった ($F(1, 24) = 0.11, p = .74, \eta^2_p = .01$)。交互作用が有意であったため, 単純主効果の検定を行った。その結果, 複数解答条件における遅延の有無の効果と, 直後条件における解答数の効果が有意であった ($ps < .02$, Figure 7)。

以上の結果から, 解答をタイピングさせた本研究においては, 遅延フィードバック条件に比べ直後フィードバック条件の方が学習が促進され, 単語対を用いた場合は遅延フィードバックでは誤検索効果が生じないという先行研究の結果 (e.g., Grimaldi & Karpicke, 2012; Hays et al., 2013; Kornell, 2014) と一致した結果が得られた。しかし, 複数遅延条件は複数直後条件に比べて成績が低かった。誤検索効果のメカニズムを説明するとされる *elaborative retrieval* 説や *search-set* 説に基づけば, 意味的なネットワークにおける情報の活性化量を増やせば遅延フィードバック条件下でも学習促進が生じるので, 遅延条件であっても複数解答させれば評価テスト成績は低下しないと考えられる。しかし研究 2 の結果はこの仮説を支持しなかった。むしろ複数遅延条件の成績が低下したのは, 注意分割によって十分な符号化がなされなかったために, 検索語や, 検索語とターゲット語の符号化が十分に行われず, 遅延の間に検索手がかりが減衰した可能性がある。一方直後フィードバックの場合, 複数解答させた方が正答率は向上した。これはタイピングにより注意分割が生じて検索語の符号化が抑制されたが, 複数解答させたことによって検索語が符号化される可能性が高まり, 結果として学習促進が生じたと考えられる。

本研究の結果より, 研究 1 で繰り返し確認された遅延フィードバック

下での誤検索効果は、先行研究との検索語の解答方法の違いによるものであり、検索語をタイピングさせると遅延フィードバック下では学習促進効果が消失する可能性が示唆された。Rapp & Fischer-Baum (2014) の示唆する通り、タイピングを課したことによって押すべきキーへの注意分割が生じたとすれば、タイピングが課されていた事前テスト時の情報の符号化が阻害され、それによって誤検索効果が消失した可能性がある。

しかしタイピング能力は前節で述べた通り個人差が大きい。そこで、続く研究 3 では単語学習課題と並行してより一般的な副課題を実施し、注意分割が誤検索効果に及ぼす影響について検討した。

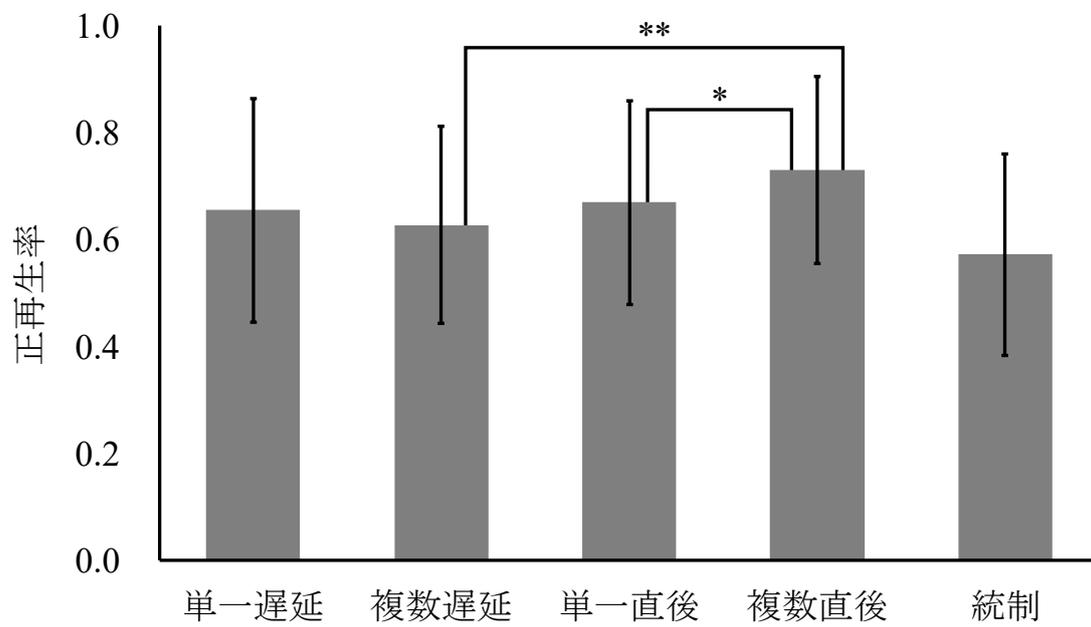


Figure 7. 研究 2 の各条件の評価テスト正答率 (** $p < .01$, * $p < .05$)

第3節 注意分割の影響（研究3）

1. 目的

ここまでの研究で、遅延フィードバック事態でも誤検索効果が生じることが繰り返し確認されたが、研究2より解答方法をタイピングにすると学習促進効果が消失する可能性が示唆された。続く研究3では Buchin & Mulligan (2017) を踏襲し、事前テストと並行してより一般的な副課題を実施し、注意分割が誤検索効果に及ぼす影響について検討した。

2. 方法

参加者 母語を日本語とする大学生、短期大学生 47 名（男性 9 名、女性 38 名、平均年齢 18.3 歳、 $SD = 0.7$ ）が実験に参加した。これらの参加者のうち 24 名（男性 5 名、女性 19 名、平均年齢 18.5 歳、 $SD = 0.9$ ）を副課題なし群、残り 23 名（男性 4 名、女性 19 名、平均年齢 18.1 歳、 $SD = 0.3$ ）を副課題あり群に振り分けた。

刺激と条件 主課題である単語学習課題には、これまでの研究と同一の単語対 100 組を用いた。すべての参加者が研究 1-1, 1-3, 研究 2 同様、検索語の解答数（2 水準）×フィードバック遅延の有無（2 水準）の 4 条件に、統制条件を加えた 5 条件で単語対を学習した。副課題（後述する数字判断課題）は、Buchin & Mulligan (2017) の実験 3 を踏襲し、1 桁の自然数（1—9）を使用した。この課題は副課題あり群の参加者のみに実施した。

手続き 研究 1-3 とほぼ同一の手続きを用いた。相違点として、事前テストでターゲット単語を推測する際、副課題あり群の参加者には数字判断課題を実施した。

副課題あり群は、学習フェーズの事前テストの検索語解答時間を 1 単

語につき 6 秒に固定したこと、事前テストと並行して副課題である数字の奇数・偶数判断課題を行わせたこと以外は研究 1-3 と同様の手続きで実施した。事前テストでは、検索条件の手がかり単語の提示と同時に、2 秒に 1 つ数字を聴覚提示した。検索語の解答を求める画面ののち、参加者には提示された数字のうち奇数の個数（0—3）をテンキーで入力させた。入力した数字が正しい場合には「正解です」の文字を提示し、誤っている場合は「違います」の文字提示と同時にビーブ音を鳴らした。複数解答条件は、以上の手続きを 2 度繰り返した。

副課題なし群は、事前テストにおける単語の解答時間、および 6 秒ごとに検索語の解答を求める画面を提示した（複数解答条件はそれを 2 度繰り返した）こと以外は研究 1-3 と同一の手続きで実施した（Figure 8, Figure 9）。

3. 結果と考察

本研究においても、検索語とターゲット単語が同一であった試行（4%）は除外して分析を行った。まず副課題の成績について、副課題あり群における副課題の平均正答率は、単一条件が .86 ($SD = .11$)、複数条件が .84 ($SD = .09$) であり、検索語の解答数（つまり副課題の回数）による成績の差はなく ($t(22) = 1.45, p = .16, d = .20$, Table 4)⁷、参加者には検索語の解答数によらず、副課題によって同様の負荷がかかっていたと考えら

⁷ 分散の正規性を満たしていない変数が含まれていたため、変数に対して逆正弦変換を行ったが、それでも正規性は満たされなかった。 t 検定は分散の正規性に対してある程度頑健であり、また副課題の成績に対する分析は本研究においては予備的なものであるため、これ以上の分析は行わなかった。

れる。

続いて本研究でも誤検索効果が生じたかどうかを検討するため、副課題の有無による検索条件（4条件の成績をプールしたもの）と統制条件の評価テスト成績を比較した。二要因分散分析を実施した結果、検索の有無の主効果と副課題の有無の主効果が有意であった（それぞれ $F(1, 45) = 45.02, p < .01, \eta^2_p = .50$; $F(1, 45) = 5.81, p = .02, \eta^2_p = .11$ ）。交互作用は有意ではなかった（ $F(1, 45) = 0.04, p = .85, \eta^2_p < .01$ ）。⁸

つまり、本研究においてもこれまで同様の誤検索効果が生じたが、副課題あり群は副課題を課していない統制条件の成績も低かった。この原因の一つとして、統制条件の刺激が提示されている間に、検索条件の刺激についての処理を行っていた可能性がある。副課題によって検索条件の刺激を時間内に十分処理できず、統制条件の刺激が提示されている間に検索条件の単語や検索語を想起していたかもしれない。それ以外の可能性のある原因については今後の課題で述べる。

次に、4つの検索条件に対して副課題が及ぼした影響について検討するため、副課題の有無（参加者間要因、2水準）×解答数（参加者内要因、2水準）×フィードバック遅延の有無（参加者内要因、2水準）の3要因分散分析を実施した（Table 5）。その結果、2次の交互作用は有意ではなかった（ $F(1, 45) = 0.49, p = .49, \eta^2_p = .01$ ）。また1次の交互作用もすべて有意ではなかった（副課題の有無×遅延の有無 $F(1, 45) = 2.23, p = .14, \eta^2_p = .05$; 副課題の有無×解答数 $F(1, 45) = 0.13, p = .72, \eta^2_p$

⁸ 逆正弦変換を行ったものに対して同様の分析を行った結果、分析結果の傾向は変わらなかった（検索の有無の主効果 $F(1, 45) = 41.10, p < .01, \eta^2_p = .48$; 副課題の有無の主効果 $F(1, 45) = 5.80, p = .02, \eta^2_p = .11$; 検索の有無と副課題の有無の交互作用 $F(1, 45) < 0.01, p = .96, \eta^2_p < .01$ ）。

< .01; 遅延の有無×解答数 $F(1, 45) < 0.01, p = .99, \eta^2_p < .01$ 。最後に、副課題の主効果と遅延の有無の主効果が有意であり（それぞれ $F(1, 45) = 6.38, p = .02, \eta^2_p = .12$; $F(1, 45) = 9.02, p < .01, \eta^2_p = .17$ ）、解答数の主効果は有意でなかった（ $F(1, 45) = 0.55, p = .47, \eta^2_p = .01$ ）。⁹

フィードバック遅延の効果については主効果のみが有意であり、一次、二次の交互作用は有意ではなかった。つまり副課題の有無によらず遅延フィードバックは誤検索効果を阻害したといえる。しかし、副課題の有無×遅延の有無の一次の交互作用にのみある程度の大きさの効果量があり、また本研究において最も検討すべき分析のため、追加として解答数をプールしたフィードバック遅延の有無要因の各水準における副課題の有無の要因の単純主効果の検定を行った。その結果、フィードバック遅延の有無によらず、副課題の単純主効果が有意であった（ $ps < .01$ ）。また副課題の有無の要因の各水準におけるフィードバックタイミング要因の単純主効果の検定を行った結果、副課題あり群でのみ、フィードバックタイミングの単純主効果が有意であった（ $p = .01$, Figure 10）。¹⁰

⁹ 逆正弦変換を行ったものに対して同様の分析を行った結果、分析結果の傾向は変わらなかった（2次の交互作用 $F(1, 45) = 0.18, p = .68, \eta^2_p < .01$; 副課題の有無×遅延の有無の交互作用 $F(1, 45) = 1.70, p = .20, \eta^2_p = .04$; 副課題の有無×解答数 $F(1, 45) = 0.23, p = .63, \eta^2_p = .01$; 遅延の有無×解答数 $F(1, 45) = 0.04, p = .84, \eta^2_p < .01$; 副課題の主効果 $F(1, 45) = 6.42, p = .02, \eta^2_p = .13$; 遅延の有無の主効果 $F(1, 45) = 7.41, p = .01, \eta^2_p = .14$; 解答数の主効果 $F(1, 45) = 1.04, p = .31, \eta^2_p = .02$ ）。

¹⁰ 逆正弦変換を行ったものに対して同様の分析を行った結果、分析結果の傾向は変わらなかった（フィードバック遅延の有無によらず、副課題の単純主効果が有意であった（ $ps < .05$ ）。同様に、副課題の有無の要因の各水準におけるフィードバックタイミング要因の単純主効果の検定を行った結果、副課題あり群でのみ、フィードバックタイミングの単純主効果が有意であった（ $p = .01$ ））。

以上の結果をまとめると、まず検索する情報量（解答数）は主効果もそのほかの要因との交互作用も有意ではなかったことから、複数の情報を検索させて意味ネットワークの活性化を促しても、誤検索効果には影響しないという研究 1 の結果が研究 3 においても再現された。続いて副課題の有無はその主効果が有意であった。これより、テスト効果における注意の影響（e.g., Buchin & Mulligan, 2017）とは異なり、誤検索効果においては事前テストでの検索中や検索語の解答時に注意を分割することが学習促進を阻害することが示された。注意分割は符号化を抑制し、検索には影響しないといわれる（*Craik et al., 1996; Mulligan, 2008*）ことから、研究 2 と同様に誤検索効果の生起には事前テストにおいて検索した情報の符号化時の処理が重要である可能性がある。

更に、追加として行った副課題の有無×遅延の有無の一次の交互作用についての単純主効果の検定結果より、副課題なし群には遅延による効果はほとんどなかった。一方の副課題あり群においては、遅延フィードバックの状況下では学習促進効果が小さくなった。これを踏まえると、*Grimaldi & Karpicke (2012)* 等の先行研究で繰り返し示されてきた遅延フィードバック条件下で誤検索効果が生じなくなる現象は、事前テスト時に何らかの注意分割が生じ、そのために本来生じるはずの誤検索による学習促進効果が小さくなり、結果として統制条件と同程度の成績になっていたためであるといえるであろう。

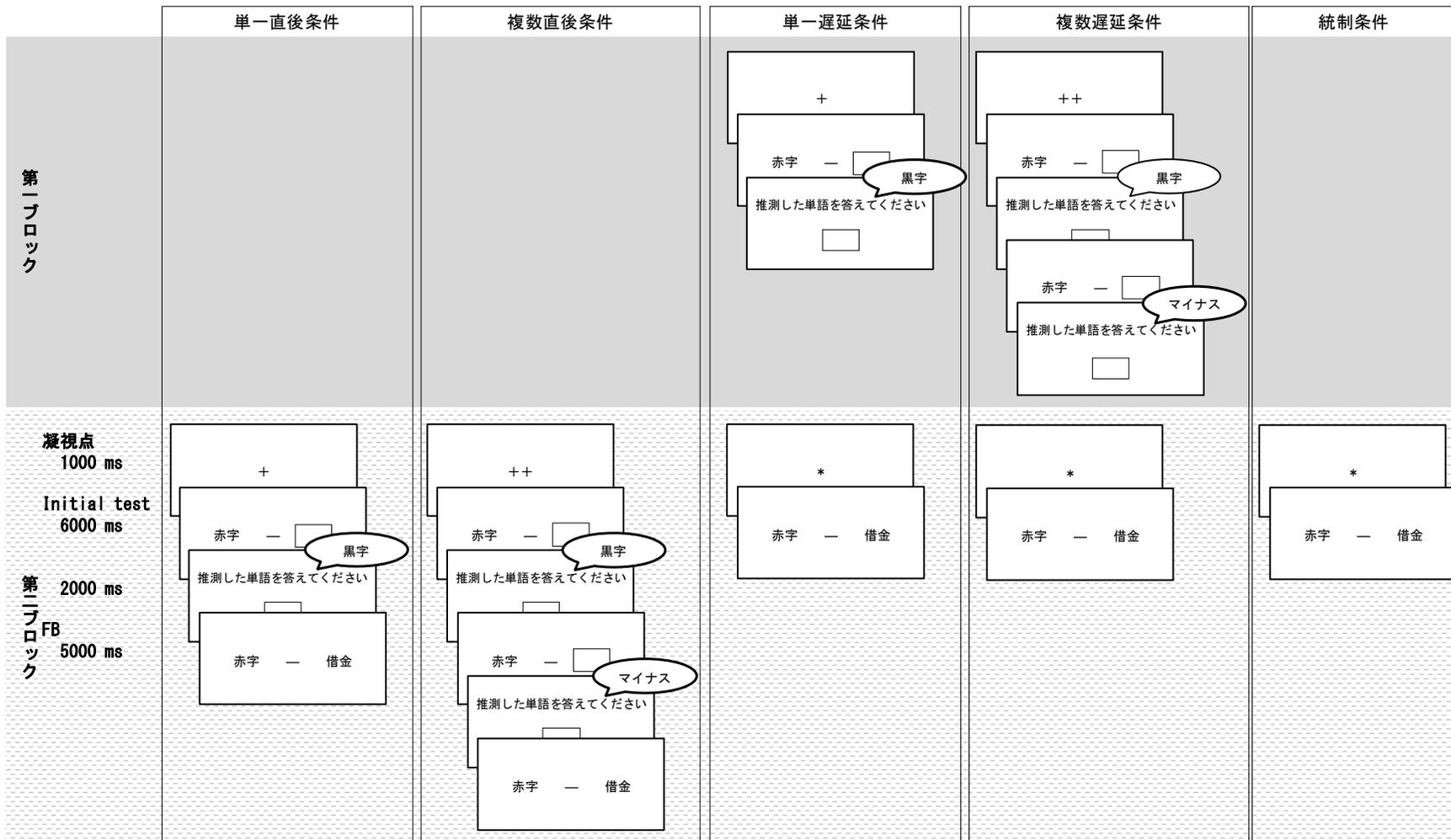
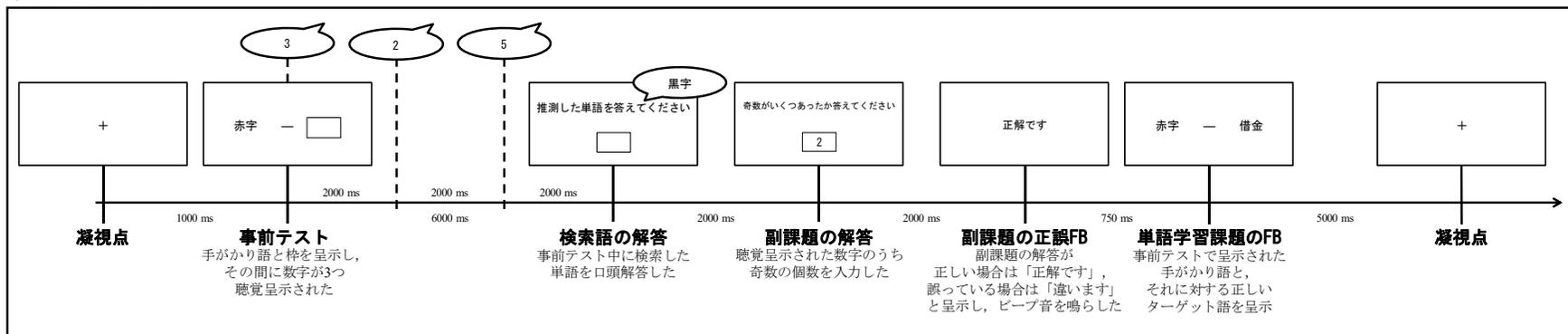


Figure 8. 研究3の副課題なし群における各条件の刺激提示順序

副課題あり群



副課題なし群

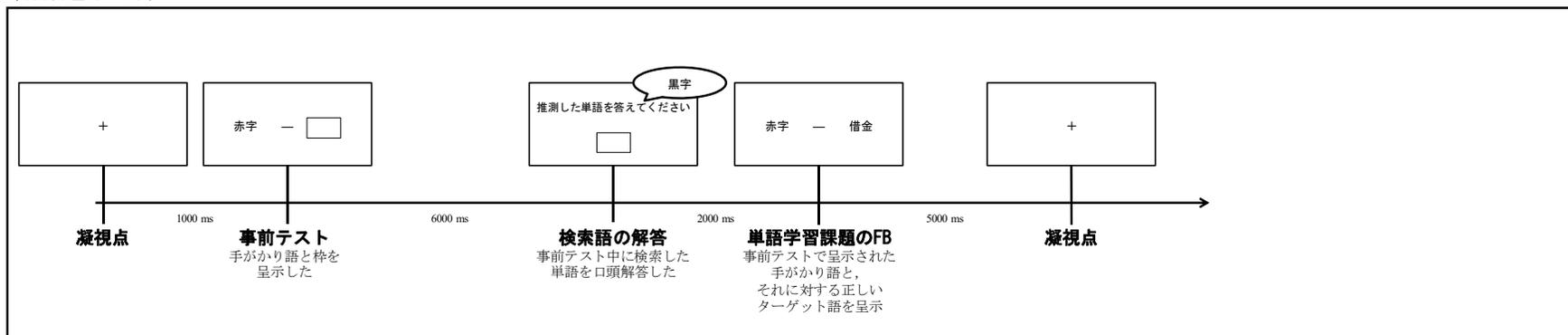


Figure 9. 研究3の副課題あり群と副課題なし群の単一直後条件の刺激提示例

Table 4

研究 3 の検索語の解答数別の，副課題の平均正答率（表中のカッコ内は *SD* を示す）

単一条件	複数条件		
	全体	1 回目	2 回目
.86 (.11)	.84 (.09)	.86 (.09)	.82 (.10)

Table 5

研究3の各条件評価テスト正答率

	平均	SD
全体		
統制条件		
単一遅延条件	.69	.19
複数遅延条件	.68	.19
単一直後条件	.73	.17
複数直後条件	.72	.15
検索条件全体		
副課題あり		
統制条件	.51	.21
単一遅延条件	.62	.21
複数遅延条件	.61	.21
単一直後条件	.68	.19
複数直後条件	.69	.13
検索条件全体	.65	.16
副課題なし		
統制条件	.63	.18
単一遅延条件	.74	.16
複数遅延条件	.74	.15
単一直後条件	.78	.13
複数直後条件	.75	.16
検索条件全体	.75	.63

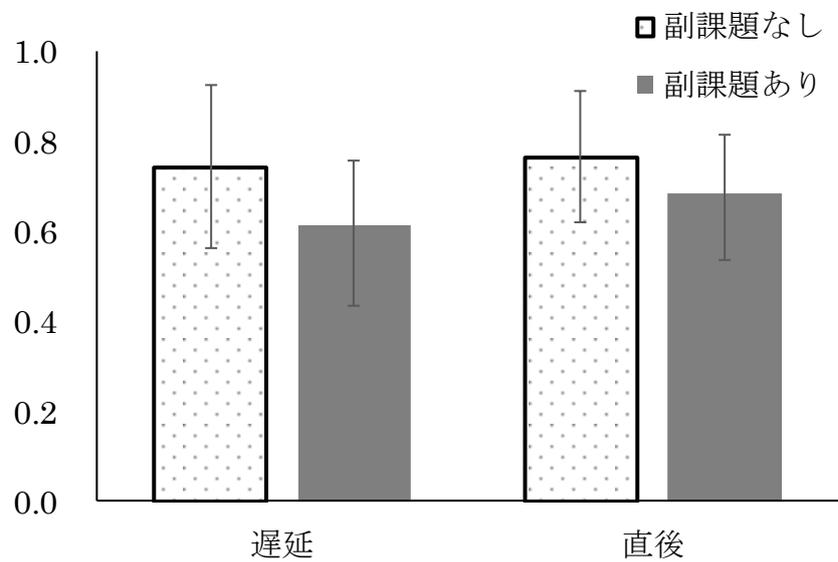


Figure 10. 研究3の検索数をプールした、フィードバックタイミングと副課題の有無による評価テスト正答率

第 3 章 総合考察

第 1 節 本研究の成果と意義

1. 本研究の成果

1.1 本研究のまとめ

本研究では、テストで誤った場合の学習促進（誤検索効果）のメカニズムを明らかにすることを目的とした。従来、誤検索効果は事前テスト時に生じる意味的なネットワークの活性化によって生じるが、特に単語対の場合その活性化は短時間しか持続しないため「刺激として単語対を用い、事前テスト後正答フィードバックをすぐに与えない」条件下においては生じないと言われてきた（Grimaldi & Karpicke, 2012; Hays et al., 2013; Kornell, 2014）。そこで意味ネットワークの活性化によって遅延フィードバック下での学習促進が生じるようになるか検討した。

研究 1 では意味ネットワークの活性化量（Grimaldi & Karpicke, 2012 や SAM model などの先行研究を踏まえ、事前テスト時に解答する検索語の数）が誤検索効果に及ぼす影響について検討した。研究 1-1 では、先行研究で繰り返し示されてきた遅延フィードバック状況下での学習促進効果の消失が、検索する誤情報の量を増やして意味的なネットワークにおいて活性化させる情報量を増やすことで防げるか検討した。その結果、意味的なネットワークにおける情報の活性化量は、フィードバック提示までの遅延の有無によらず評価テスト成績に影響しなかった。また、そもそも先行研究とは異なり、単一遅延条件であっても（つまり、先行研究で繰り返し学習促進が生じないと言われてきた、誤った検索語を 1 つだけ解答し、数分後に正答フィードバックを提示する条件下でも）誤検索効果が生じた。

続く研究 1-2 では研究 1-1 で生じていた遅延時間のばらつきを統制し、

遅延時間と誤検索効果の関係について検討した。その結果、遅延時間の有無・長短によらず、いずれの条件下でも誤検索効果が生じた。しかし統計的には有意でないものの、遅延時間が長くなるほど評価テストの成績が低下する傾向がみられた。

この結果を踏まえ、研究 1-3 では学習フェーズを 2 つのブロックに分け、遅延時間を 15 分から 25 分に延長した場合に、意味的なネットワークにおいて活性化させる情報量が誤検索効果に影響するかどうか検討した。その結果、これまでの研究同様、意味的な情報の活性化量は誤検索効果に影響せず、また正答フィードバック提示まで先行研究より長い遅延時間を設けても誤検索効果による学習促進効果は消失しなかった。

以上の研究 1 の結果より、誤情報検索時に活性化する意味的なネットワークの情報量を増やすことは誤検索効果に影響しない可能性が示された。しかし、先行研究で繰り返し再現されてきた「遅延フィードバック下での誤検索効果の消失」も見られなかった。この理由の一つとして検索語の解答方法が考えられたため、研究 2 では、研究 1-3 と同様の手続きを用いて、検索語をタイピングで解答させた。その結果、遅延条件は直後条件に比べて再生成績が低下するという先行研究の結果が再現された。この結果と先行研究が解答方法としてタイピングを用いていることを踏まえると、検索語の解答方法が誤検索効果、特に遅延フィードバックでの誤検索効果に影響していると考えられる。

第 2 章第 2 節で述べた通り、口頭で解答することは符号化を促進し (Forrin et al., 2012)、またタイピング時にはワーキングメモリを必要とするような処理が要求されることから (Pinet & Nozari, 2018; Rapp & Fischer-Baum, 2014)、注意分割によって符号化が抑制された可能性がある (Craik et al., 1996; Mulligan, 2008)。そこで研究 3 では単語学習課題と

並行して Buchin & Mulligan (2017) と同様の副課題 (数字判断課題) を実施し, 副課題による注意分割が誤検索効果に及ぼす影響について検討した。その結果, 副課題を課した群の成績が低下した。また副課題あり群でのみ, 遅延フィードバック時に成績が低下した。つまり副課題によって注意分割が生じ, 副課題と並行して入力された情報の符号化が抑制されたことが誤検索効果に影響したと考えられる。加えて, 研究3でも解答数は評価テスト成績に影響しなかった。以上を総合すると, 検索語を口頭解答させた研究1は検索語の符号化が促進され, タイピングさせた先行研究や研究2は副課題を課した研究3同様検索語が十分符号化されなかったために遅延フィードバック時に成績が低下した可能性がある。

1.2 誤検索効果のメカニズム

Search-set 説や, その元となった elaborative retrieval 説が仮定する通り, 誤検索効果が意味的なネットワークの豊かさ (本研究における検索語の解答数) によるものであるとすれば, 複数解答させることは意味的な精緻化を促し (cf. Lehman & Karpicke, 2016), その結果直後条件も遅延条件も成績が向上すると考えられる (ただし天井効果が生じ, 直後条件では成績向上は見られない可能性がある)。しかし, 検索語の数の効果は研究2を除いて評価テスト成績に影響しなかった。研究2においても, 直後条件では単一条件より複数条件の成績が高かったが, 複数解答することがより学習を促進させると考えられる遅延条件ではむしろ成績を低下させた。では, search-set 説, elaborative retrieval 説の観点からは, 本研究の結果はどのように説明できるであろうか。これらの説は, 刺激が有する意味ネットワークの活性化によって想起成績が向上するという点で一致しており, 特に elaborative retrieval 説はそのような意味的なネット

ワークに基づいて有効な検索手がかりが形成されることが想起成績を向上すると仮定する。そこで、まず search-set 説, elaborative retrieval 説の観点からの誤検索効果の生起メカニズムを整理してみよう。

単語対の学習は、手がかり単語とターゲット単語のそれぞれの符号化と対の符号化（単語間の連合の形成、強化）とによってなされる。誤検索効果のパラダイムでは、事前テストにおいて手がかり単語のみが提示され、ターゲット単語（本研究における「検索語」）の自己生成が求められる。この際に、意味ネットワークにおける検索が生じ、手がかり単語と強い連想価を持つ単語が検索語として生成され、手がかり単語と検索語との対の表象が形成されると考えられる。この後のフィードバックでは、正しいターゲット単語として比較的連想価の低い単語が手がかり単語と対で提示され、参加者は手がかり単語とターゲット単語との対連合学習を行う。この時、事前テスト時に形成された手がかり単語と自己生成した検索語との対に加えて手がかり単語とターゲット単語、あるいは検索語とターゲット単語との対連合も形成されると考えられる。これによって、事前テスト時に生成された検索語は、評価テストにおけるターゲット単語の検索手がかりとなる。評価テスト時に手がかり単語が提示されると、それと連想価と高い検索手がかりが活性化し、正答（ターゲット単語）へのアクセスが促進される。

このようなメカニズムを想定すると、本研究の結果は以下のように予測される。事前テストにおいて活性化させる情報量が増加すると意味的なネットワークがより活性化し、更に検索手がかりも増えるため、評価テストの成績が増加し、遅延しても誤検索効果が減衰しにくいであろう。

しかし、結果は予測に反し、単語対の学習における誤検索効果は事前テストの回数や遅延の影響を受けないというものであった（研究 1）。こ

の結果を上記のメカニズムで解釈すると、1回の事前テストで付与される検索手がかりの有効性で説明できると考えられる。1回目の事前テストで自己生成される検索語は、手がかり単語と強い連想関係もつ単語である可能性が高い。そのため、手がかり単語を介して正しいターゲット単語の検索手がかりとして有効に機能するものであり、また本研究で設定した遅延時間ではその連合の活性化が減衰しにくいものであった可能性が考えられる。

また、1回目の事前テストにおいて、手がかり単語からターゲット単語を検索する過程で連想強度の高い複数の候補へのアクセスが生じていた可能性も考えられる。そのうちの1つが検索語として自己生成されるが、他の候補の活性値も上昇しており、それらが検索手がかりとして機能したことも考えられる。2回目の事前テストでは、1回目の事前テストでネットワークにおいて活性値の上昇した範囲から2個目の検索語を選択しているとする、単純に情報量が2倍になったとは言えない可能性もある。つまり、事前テスト1回で十分な量の情報が活性化し、有効な検索手がかりが形成され、誤検索効果が生じていたという解釈である。

以上の解釈は、事前テストにおける検索手がかりの形成の過程に十分な処理資源が充当されていることが前提となることが本研究からは示唆される。すなわち、研究2、3の結果と上記の解釈の統合という点である。

研究2では、事前テストにおける解答方法を研究1での口頭解答からタイピングに変更した。この手続きの変更によって、事前テスト時に含まれる検索手がかり形成の過程のうち手がかり単語から検索語を検索する過程は影響を受けないものの、自己生成された検索語の符号化や手がかり単語と検索語との連合の符号化が抑制されたと考えられる。その結

果として、1 回目の事前テストのみでは十分に有効な検索手がかりが形成されなかったと考えられる。2 回目の事前テストによって 2 個目の検索手がかりが作られることによって、正しいターゲット単語の検索を促進するのに十分な検索手がかりが形成され、誤検索効果が生じたという解釈である。しかし、この複数の検索手がかりによるターゲット単語の検索の促進は直後条件でしか生じていない。このこともまた、符号化が抑制されたために十分な強度を持つ単語の記憶や対連合が形成されていなかったことを示唆している。

一方研究 3 では、副課題を課すことによって事前テスト遂行時の全過程に負荷を加えたと考えられる。すなわち、手がかり単語からの検索語の検索、検索語の保持、検索語の符号化、検索語と手がかり単語との連合の符号化に十分な処理資源が充当されない事態である。これにより、事前テストにおいて有効な検索手がかりが形成できなかった可能性がある。

副課題のない状況下では、同様の条件で行われた研究 1-1 および研究 1-3 の結果が再現されているが、事前テスト遂行時に負荷を加えられた副課題あり条件では、事前テストの回数の効果は見られず、遅延の効果のみが確認されている。この結果は、事前テストを行うことだけでなく、事前テスト遂行時に十分な処理が行われ、有効な検索手がかりが形成されることが誤検索効果の生起に重要であることを示唆するものと言えるであろう。

以上をまとめると、事前テストにおいて活性化される情報量は誤検索効果に対して影響するものの、単語対の学習では、1 回の事前テストで形成される検索手がかりが強力であるために、研究 1 では 2 回目の事前テストによって追加される情報の影響が隠蔽されていたという解釈を行う

ことによって、本研究の結果が意味的なネットワークの活性化を想定する説、特に検索語がターゲット単語を想起する際の媒介として機能すると想定する *elaborative retrieval* 説で一貫して説明できる可能性があると言えるであろう。

本研究では、研究 1 においては遅延の長短による量の効果は検証できなかった。しかし、研究 2 においてタイピングでの解答を課し、1 回目の事前テストにおける検索手がかり形成に干渉する状況を設定したところ、遅延の効果、検索する情報量の効果がともに確認された。このことから、さらに長い遅延を置くなどの改変を加えることによって 1 回目の事前テストで形成される検索手がかりの有効性を減じることができるならば、研究 1 の手続きにおいても遅延の効果、量の効果が検証できる可能性が示唆される。

本研究の結果から、誤検索効果は「事前テストにおいて記録すべき正答とは異なる情報（誤情報）が符号化され、それが評価テスト時に有効な検索手がかりとなることによって学習促進が生じる」と結論できる（Figure 11）。直後フィードバックの条件下での誤検索効果においては、評価テスト時に事前テストでの誤答（本研究における検索語に相当するもの）を想起できる項目の方が、事前テストの誤答を想起できない項目よりも評価テスト成績が高い、つまり正答と事前テストでの誤答との対連合を示す研究もある（e.g., Yan, Yu, Garcia, & Bjork, 2014）。

加えて、テスト効果を説明する *episodic context* 説は「学習時と検索時の文脈の変化」によって学習促進を説明する理論であるため、今回の結果そのものを説明するものではないが、記録すべき情報とは異なる情報の符号化によって学習促進を説明しようとしている点では、本研究の結果と部分的ではあるが合致する。先の段落で、本研究において複数解答

することが誤検索効果を促進しなかったのは、1回の事前テストで生成された検索語が検索手がかりとして十分に有効であった可能性を挙げた。しかし2回の事前テストが連続して実施されたため、「評価テスト時に手がかりとなるような、複数の文脈が統合された有効な学習時の文脈を提供しなかった」とも考えられる。事前テストを分散して実施した場合に評価テスト成績が向上すれば、事前テスト時の文脈もフィードバックの符号化に貢献していることが示唆できるであろう。

ただし、意味的な情報が活性化しにくいと考えられる未知の語（未知の外国語や、使用頻度の著しく低い母国語の難語）とその定義の対連合学習でも誤検索効果が生じることから、意味的な情報は誤検索効果の生起には重要でないという指摘もある（Potts & Shanks, 2014）。しかしこのような単語対を用いた誤検索効果においても、事前テストで解答した誤情報が手がかりとターゲットをつなぐ媒介として機能している可能性があり、この点については今後さらに検討する必要があるであろう。

2. 本研究の意義

学習者にとって誤りは基本的にネガティブなイベントであり、頻発することは学習意欲の低下にもつながるが、学習成績の低い者ほどその頻度は高く、また学習時の誤りをなくすことは難しい。しかし本研究の結果から、例えばテストに集中して取り組むことは、たとえその結果が芳しくなくても、テキストを読むこと以上の学習促進効果があると考えられる。またテストの答え合わせ（本研究におけるフィードバック）は、可能な限り早い方が学習促進効果は保証されるが、誤りに注意を向け符号化を促すことで遅延フィードバックでも学習を促進すると考えられる。

本研究で扱った刺激は単語対であるため、後述の通り学校教育の場に

おけるテストの効果に本研究の結果をそのまま応用することは難しいが、テストにおける誤りへ注意を向けること（符号化）が後の学習を促進するという結果は、今後誤検索効果をはじめとする「誤り」を活用する学習方略を実際の学習場面に応用する際に重要な示唆となり得る。テスト自体は学校教育の場において広く使われていることから、テストでの正答はもちろん、誤りを学習促進に有効に活用する方法やそのメカニズムが今後さらに明らかになれば、教授者はテストにおける誤りを用いてより良い指導ができ、同時に学習者自身も自身の誤りをより活用できるようになり、ひいては自己教育力の育成にもつながるであろう。

第2節 今後の課題

最後に、本研究の今後の課題として以下の4点を挙げる。第一に使用した刺激がすべて単語対であるため、文章のようなより複雑な刺激を用いた場合に、特に注意の影響については明らかでない。本研究において単語対を使用したのは、誤検索効果のメカニズムを明らかにするために剰余変数（手がかりに対する参加者の知識量や興味など）を統制することが目的であった。しかし学校教育の場に応用するためには、学校などで実際に使用されている刺激を用いながらさらに検討していく必要がある。また、先行研究で主に使用されている刺激がアルファベットの単語であることから、本研究との結果の違いが刺激の表記によるものではないことを保障するような追試も必要であろう。さらに現在の教育においては「誤り」などの失敗を肯定的にとらえ、成功に結び付けることが推奨されている（西村・瀬尾・植阪・マナロ・田中・市川，2017）。西村ら（2017）は学業場面での失敗観が失敗の活用可能性の認知と失敗の脅威

性の認知によって構成されていることを示している。本研究の知見は失敗の活用可能性の指導に生かせるであろう。

第二に、本研究では意味的なネットワークにおいて活性化される情報量を増やすための操作として事前テストでの検索語の解答数を操作したが、解答数をさらに増やすなどより活性化するような操作を加えた場合に意味的なネットワークの豊かさが誤検索効果に影響する可能性は否定できない。複数解答させる操作が意味的なネットワークを十分には豊かにしなかったとしても、**search-set** 説や **elaborative retrieval** 説に基づけば記録すべき情報と関連する情報が複数あることは誤検索効果を促進するとも考えられるが、この点については今後さらに検討する必要があるであろう。

第三に、先行研究を踏まえ注意分割によって情報の符号化が妨害された可能性を挙げたが、本研究のパラダイムにおいて実際に事前テスト時の情報（例えば参加者が誤って答えた検索語）の符号化が妨害されているかは直接検討されていない。今後は評価テスト時にターゲットのみを再生させるのではなく、事前テスト時の誤答（検索語）も解答させるような実験を実施してより詳細な誤検索効果のメカニズムを明らかにする必要がある。これと関連して、研究3において、副課題あり群は課題を課していない統制条件の評価成績も低かった。このような結果が得られた原因として、第2章第3節で統制条件中に検索条件の刺激を処理していた可能性を挙げたが、それ以外にも副課題あり群となし群の参加者が等質でなかった可能性がある。実験実施時に参加者を各群にランダムに割り当てたが、偶然偏りが生じた可能性は否定できない。研究3では研究1, 2の手続きを踏襲するため、また単語対の学習条件が5条件あり、副課題の有無を参加者内で操作すると各学習条件の刺激数が少なくなり

すぎるために、副課題の有無を参加者間要因としたが、これらの可能性について検討するためには、今後副課題の有無を参加者内で操作するなどの手続きの改善を行って再検討する必要があるであろう。

最後に参加者の要因が挙げられる。本研究の参加者はすべて大学生・短期大学生であり、今回誤検索の効果を生じた手続きがより若い子どもたちに対しても同様の結果を生じるかどうかは明らかでない。学校教育の場における誤検索効果について検討するためには、小学生や中学生など実際に応用を狙う世代を対象として効果を検証していく必要があるであろう。また、本論文を構成するすべての研究において、男性よりも女性の方が多かった。先行研究では参加者の男女の内訳について言及していないものも多く、本研究の結果に必ずしも性差が影響しているとは言えないが、より若い子どもたちに対して実施する場合には性別など本研究で扱わなかった特性も統制した実験を行う必要があるであろう。

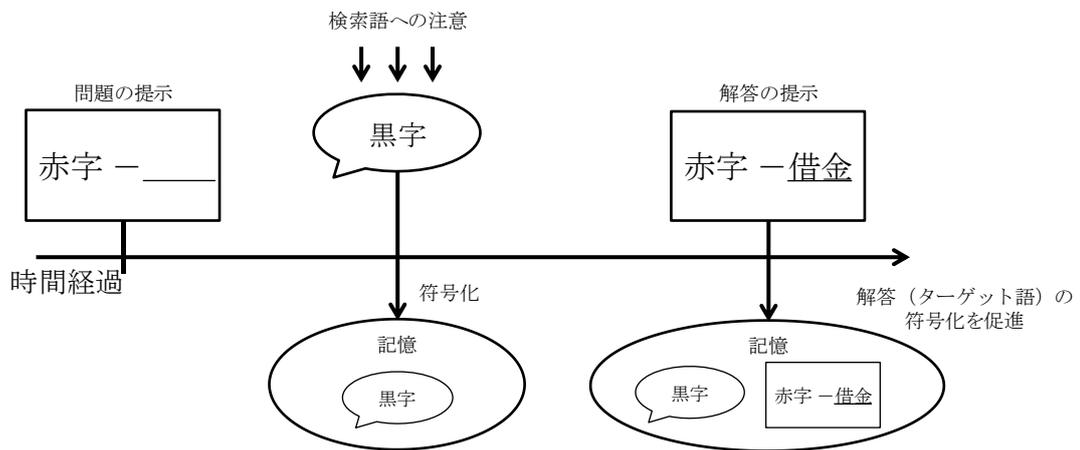


Figure 11. 本研究から示唆される誤検索効果のメカニズム

引用文献

- Bacon, F., (1870–1876). *The works of Francis Bacon* (J. Spedding, R. L. Ellis, & D. D. Heath, Col. & Ed). (Vols. 1, 4–5) London: Longmans & Co. (Original work published 1620)
- (ベーコン, F. 服部英次郎 (訳) (2005). ノヴム・オルガヌム 服部英次郎・多田英次 (他訳) 世界の大思想 II-4 ベーコン 学問の進歩／ノヴム・オルガヌム他 ワイド版 (pp. 199–411) 河出書房新社)
- Baddeley, A., & Wilson, B. A., (1994). When implicit learning fails: Amnesia and problem of error elimination. *Neuropsychologia*, 32, 53-68.
- 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.
- Bridger, E. K., & Mecklinger, A. (2014). Errorful and errorless learning: The impact of cue-target constraint in learning from errors. *Memory & Cognition*, 42, 898-911.
- Buchin, Z. L., & Mulligan, N. W. (2017). The testing effect under divided attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 1934–1947.
- Butler, A. C., & Roediger, H. L. III. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, 19, 514–527.
- Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 35, 1563–1569.

- Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *37*, 1547–1552.
- Carpenter, S. K., & DeLosh, E. L. (2005). Application of the testing and spacing effects to name-learning. *Applied Cognitive Psychology*, *19*, 619–636.
- Carpenter, S. K., & Kelly, J. W. (2012). Tests enhance retention and transfer of spatial learning. *Psychonomic Bulletin & Review*, *19*, 443–448.
- Clare, L., & Jones, R. S. P. (2008). Errorless learning in the rehabilitation of memory impairment: A critical review. *Neuropsychological Review*, *18*, 1–23.
- Coppens, L. C., Verkoeijen, P. P. J. L., & Rikers, R. M. J. P. (2011). Learning Adinkra symbols: The effect of testing. *Journal of Cognitive Psychology*, *23*, 351–357.
- Craik, F. I., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, *125*, 159–180.
- Cyr, A. A., & Anderson, N. D. (2015). Mistakes as stepping stones: Effects of errors on episodic memory among younger and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 841–850.
- 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.
- Forrin, N. D., MacLeod, C. M., & Ozubko, J. D. (2012). Widening the boundaries of the production effect. *Memory & Cognition*, *40*, 1046–

1055.

- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*, 1–67.
- Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition*, *40*, 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.
- Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*, 269–299.
- 岩田一男 (2015). 初年次教育におけるタイピング練習とその関連性についての調査研究 日本情報経営学会誌 36, 74–85.
- Izawa, C. (1969). Comparison of reinforcement and test trials in paired-associate learning. *Journal of Experimental Psychology*, *81*, 600–603.
- Izawa, C. (1970). Optimal potentiating effects and forgetting-prevention effects of tests in paired-associate learning. *Journal of Experimental Psychology*, *83*, 340–344.
- Kang, S. H. K., McDermott, K. B., & Roediger, H. L. III. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, *19*, 528–558.
- Karpicke, J. D. (2017). Retrieval-based learning: A decade of progress. In J. T. Wixted (Ed.), *Cognitive psychology of memory, Vol. 2 of Learning and memory: A comprehensive reference* (J. H. Byrne, Series Ed.) (pp. 487–514). Oxford, England: Oxford Academic Press.
- Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning:

- An episodic context account. In B. H. Ross (Ed.), *Psychology of Learning and Motivation, Vol. 61* (pp. 237–284). San Diego, CA: Academic Press.
- Kornell, N. (2014). Attempting to answer a meaningful question enhances subsequent learning even when feedback is delayed. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 106–114.
- 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.
- Kornell, N., Klein, P. J., & Rawson, K. A. (2015). Retrieval attempts enhance learning, but retrieval success (versus failure) does not matter. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*, 283–294.
- Kornell, N., & Vaughn, K. E. (2016). How retrieval attempts affect learning: A review and synthesis. *Psychology of Learning and Motivation, 65*, 183–215.
- Lehman, M., & Karpicke, J. D. (2016). Elaborative retrieval: do semantic mediators improve memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 1573–1591.
- MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 671–685.
- Marsh, E. J., Fazio, L. K., & Goswick, A. E. (2012). Memorial consequences of testing school-aged children. *Memory, 20*, 899–906.
- 三村 将 (1998). 記憶障害のリハビリテーション —間違っただ方がおぼえやすいか？ 努力した方がおぼえやすいか？— 失語症研究, *18*, 136–145.

- 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
- Mulligan, N. W. (2008). Attention and memory. *Learning and Memory: A Comprehensive Reference*, 2, 7–22.
- 中野和光 (2007). 特集 「人間力」と学力の関係を問う <9> 学力テストの結果の利用の方法 学校教育研究, 22, 121–132.
- 西村多久磨・瀬尾美紀子・植阪友理・マナロ エマニュエル・田中瑛津子・市川伸一 (2017). 学業場面に対する失敗観尺度の作成 教育心理学研究, 65, 197–210.
- Pashler, H., Zarow, G., & Triplett, B. (2003). Is temporal spacing of tests helpful even when it inflates error rates? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1051–1057.
- Pinet, S., & Nozari, N. (2018). “Twisting fingers”: the case for interactivity in typed language production. *Psychonomic Bulletin & Review*, 25, 1449–1457.
- Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General*, 143, 644–667.
- Putnam, A. L., & Roediger, H. L. (2013). Does response mode affect amount recalled or the magnitude of the testing effect? *Memory & Cognition*, 41, 36–48.
- Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, 88, 93–134.
- Rapp, B., & Ficher-Baum, S. (2014). Representation of orthographic knowledge. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production*. (pp. 338–357). New York:

Oxford University Press.

- Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied*, *15*, 243–257.
- Roediger, H. L., & Karpicke, J. D. (2006a). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*, 249–255.
- Roediger, H. L., & Karpicke, J. D. (2006b). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*, 181–210.
- Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation: Vol. 55. The psychology of learning and motivation: Cognition in education* (pp. 1–36). San Diego, CA, US: Academic Press.
- Roediger, H. L., & Pyc, M. A. (2012). Inexpensive techniques to improve education: applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition*, *1*, 242–248.
- Skinner, B. F. (1968). *The technology of teaching*. New York: Appleton-Century-Crofts.
- (スキナー, B. F. 村井実・沼野一男 (監訳) (1969). 教授工学, 東洋館出版社)
- 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.
- Vaughn, K. E., Hausman, H., & Kornell, N. (2017). Retrieval attempts enhance

learning regardless of time spent trying to retrieve. *Memory*, 25, 298–316.

Yan, V. X., Yu, Y., Garcia, M. A., & Bjork, R. A. (2014). Why does guessing incorrectly enhance, rather than impair, retention? *Memory & Cognition*. 42, 1373–1383.

Appendix

Table A

本研究で使した単語対

手がかり単語	ターゲット単語	連想強度	第一位連想語	第一位連想語の 連想強度	ターゲット単語までの 連想語数
赤字	借金	.054	黒字	.185	3
移民	難民	.054	アメリカ	.115	2
汚染	川	.054	環境	.281	3
親子	関係	.051	家族	.147	4
会議	討論	.042	会社	.185	3
価格	商品	.054	高騰	.166	3
科学	進歩	.051	理科	.169	2
火災	炎	.051	火	.179	4
家庭	円満	.051	家族	.288	3
看護	医療	.048	病院	.358	3
管理	かぎ	.048	管理人	.058	3
機械	精密	.051	工場	.144	3
苦勞	労働	.054	仕事	.163	3
検査	健康	.051	病院	.160	2
権利	国民	.048	義務	.118	4
古典	昔	.051	国語	.243	2
資格	英検	.048	試験	.093	5
刺激	感覚	.042	痛み	.086	3
資源	有限	.051	ごみ	.134	3
思考	停止	.042	考え	.137	4
自信	喪失	.048	過剰	.128	1
自然	森林	.051	森	.198	5

時代	戦国	.054	歴史	.144	2
深夜	徘徊	.042	夜	.134	3
心理	学問	.054	こころ	.182	3
世界	国	.051	地球	.185	2
危険	安全	.042	がけ	.089	2
知識	頭脳	.042	勉強	.096	5
手紙	メール	.051	紙	.083	4
電話	会話	.051	けいたい	.319	2
道具	工具	.051	便利	.077	3
道路	交通	.048	車	.249	4
帽子	太陽	.042	頭	.137	5
野菜	果物	.048	にんじん	.179	4
世論	世間	.048	調査	.163	2
論理	理屈	.042	理論	.058	2
いかだ	遭難	.047	海	.241	5
いのち	人間	.053	大切	.144	4
うどん	きつね	.053	麵	.141	5
えくぼ	頬	.047	笑顔	.381	2
おかず	弁当	.044	ごはん	.438	2
おとこ	筋肉	.050	おんな	.353	1
おなか	背中	.047	腹痛	.100	5
おむつ	子供	.053	あかちゃん	.622	2
かかと	くつ	.050	あし	.575	1
かたち	積み木	.044	まる	.125	3
かばん	通学	.050	学校	.156	4
さかな	うろこ	.053	海	.372	1

せりふ	演劇	.053	だいほん	.247	5
たばこ	害	.050	けむり	.281	1
だるま	選挙	.053	赤	.163	2
だんご	くし	.047	三兄弟	.156	4
たんす	ひきだし	.047	服	.256	6
つくえ	教室	.050	勉強	.284	4
とうふ	豆	.053	白	.306	2
となり	友達	.047	トトロ	.234	5
にきび	青春	.044	顔	.266	3
のれん	布	.050	みせ	.322	2
はかま	着物	.047	成人式	.222	4
はさみ	刃物	.053	紙	.263	4
はんこ	名前	.053	印鑑	.247	4
ふすま	たたみ	.047	和室	.153	4
ふとん	羽毛	.053	睡眠	.203	2
まつり	花火	.047	夏	.363	1
まぶた	瞳	.053	め	.550	1
りんご	青森	.050	赤	.341	2
わいろ	悪	.047	かね	.313	3
アルミ	銀色	.050	缶	.206	4
カエル	雨	.053	みどり	.299	4
カルタ	遊び	.047	正月	.545	1
ゲスト	招待	.050	客	.226	3
コイン	スロット	.053	かね	.369	3
コラム	本	.053	新聞	.266	2
サウナ	熱	.047	温泉	.136	4

シルク	高級	.053	絹	.259	3
スパイ	敵	.043	映画	.126	2
スリル	映画	.053	恐怖	.163	4
チーズ	ピザ	.047	ねずみ	.239	3
テスト	期末	.053	試験	.209	3
テニス	運動	.047	ボール	.179	5
テラス	家	.047	庭	.123	4
ドレス	女性	.047	パーティ	.189	4
ネオン	夜景	.050	光	.179	4
バイク	タイヤ	.053	乗り物	.133	4
バイト	時給	.047	かね	.279	4
パイプ	水道	.050	鉄	.140	5
バケツ	リレー	.047	みず	.488	3
パズル	難解	.043	ピース	.143	2
ビール	大人	.050	さけ	.216	4
ビデオ	映像	.047	録画	.133	4
ポンプ	空気	.050	みず	.498	1
マイク	音声	.053	カラオケ	.193	4
ミイラ	死体	.047	エジプト	.392	3
ミシン	ぬいもの	.053	裁縫	.209	4
モグラ	地中	.053	土	.389	3
モデル	ファッション	.043	雑誌	.186	4
モラル	マナー	.050	常識	.199	2
ラジオ	エフエム	.053	電波	.110	5
ランチ	ご飯	.050	ひる	.439	2
ワイン	フランス	.047	赤	.379	3

Note. 研究では「手がかり単語」と「ターゲット単語」をペアにして使用した（水野, 2009 をもとに作成）。「連想強度」は、提示された手がかり単語から連想される語を答える課題において、そのターゲット単語を 1 つ目の回答として答えた者の割合を示す。本研究では連想強度が.042—.054 となるような手がかり単語とターゲット単語を使用した。「第一位連想語」は手がかりから最も多く連想された語を示し、「第一位連想語の連想強度」は手がかり単語から連想される 1 つ目の語として第一位連想語を回答した（例えば、「赤字」から「黒字」を連想し、回答した）者の割合を示す。「ターゲット単語までの連想語数」は、その手がかり単語の連想語のうちターゲット単語よりも連想強度の高い語の数を示す。

謝辞

本論文を完成させるにあたり、本当に多くの方々のご支援とご指導を賜りました。特に指導教員である宮谷真人先生には、学部に入學し、先生のゼミに所属させていただいてから今日まで、研究についてのご指導はもちろん、研究することの楽しさや意義など多くのことを教えていただきました。ここまでの成果を博士論文としてまとめることができたのも、また今の私があるのも、先生にご指導いただけたからに他なりません。本当にありがとうございました。

また本論文の作成につきましては、指導グループの中條和光先生、湯澤正通先生、中尾敬先生をはじめ心理学講座の先生方から丁寧なご指導、重要なお指摘を多数頂戴いたしました。さらに岩手大学の岩木信喜先生からも多くのご助言をいただきました。厚くお礼を申し上げます。

最後になりましたが、研究にご協力いただいたみなさま、多くの面で支えていただいた認知心理学研究室をはじめとする心理学講座の先輩・後輩の皆様、徳島文理大学の諸先生方、そして家族に、心から感謝いたします。