

ブロック練習と交互練習の単独効果と複合効果の比較検討

—学習内容の定着度, 及び, 確信度判断の正確性に着目して—

尾之上 高 哉* 井 口 豊**

本研究では, 学習法としてのブロック練習と交互練習に注目し, それぞれの単独効果, 及び, それらを組み合わせた時の複合効果を比較検討する実験を行った。大学生 66 名が, 学習を 1 週間の間隔をあけて 2 回行い, 2 回目の学習の 1 週間後にテストを受けた。66 名は, 2 回の学習の方法が異なる次の 4 つの条件, つまり, 条件 1 (2 回ともブロック練習で学習する), 条件 2 (1 回目をブロック練習で 2 回目は交互練習で学習する), 条件 3 (1 回目を交互練習で 2 回目はブロック練習で学習する), 条件 4 (2 回とも交互練習で学習する), のいずれかに割り当てられた。条件間でテストの正答率に差があるかを分析した結果, 正答率は, 条件 4, 条件 2 及び 3, 条件 1, の順で高く (条件 4 > 2 = 3 > 1), この 3 者の間には有意差が確認された。つまり, 交互練習の機会が増えるに従って, 学習内容の定着が進むことが示された。また, 実験参加者には, テスト時に自身が想起した解答がどの程度正しいと思うかについての確信度判断を, 多段階評定を用いて行ってもらった。その確信度判断と実際のテストの得点の関連を条件毎に分析した結果, 学習者は, 定着効果を持つ学習法で学習した時の方が, 確信度判断を正確に行える可能性があることが示唆された。

キーワード: 学習法, ブロック練習, 交互練習, テストの得点, 確信度判断の正確性

問題と目的

教育で目指される目標の 1 つに, 基礎的・基本的な内容の定着, つまり, 「学習者が, 学習した内容を記憶し, 必要な時に, 適切に使用できるようにすること」がある (Kang, 2016a)。例えば, 2017 年に告示された新学習指導要領では, 基礎的・基本的な内容の定着を確実に図ること, という種の文言がたびたび登場する (文部科学省, 2017)。学術研究では, 基礎的・基本的な内容の定着に効果のある学習法に関する議論が, 現在も盛んに行われている (e.g., Weinstein, Madan, & Sumeracki, 2018)。基礎的・基本的な内容の定着は, それ自身が重要な意味を持つだけでなく, 問題解決や推論といった高次の学習を支える役割もある (Kang, 2016a) ため, 教育で重要視されるのだと思われる。

基礎的・基本的な内容の定着に向けた議論の中で, 近年の米国では, 「算数・数学科の学習を, ブロック練習 (blocked practice) と交互練習 (interleaved practice) という観点から捉える」研究パラダイム (e.g., Kang, 2016b) が, 1 つの枠組みとして受け入れられつつあるようである。そのことは, この枠組みに基づく実証研究の

数自体は未だ少ないものの (Booth et al., 2017; Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013), その考え方が, 基礎的・基本的な内容の定着に効果のある学習法を議論する論文 (e.g., Weinstein et al., 2018) や, 教師や学習者向けの図書 (e.g., Busch & Watson, 2019) において, たびたび取り上げられていることからうかがえる。

ブロック練習と交互練習は, 複数の練習問題を解き学習する際に, 「問題解決に必要な方略, 或いは, 問題解決に必要な知識 (以下, 方略と呼ぶ)」に関して, 同一のものを連続的に使用するか否か, という点で区別される (Rohrer, 2009; Weinstein et al., 2018)。つまり, ブロック練習では, ある特定の方略を連続的に使用する。例えば, 方略 a を使う問題を, 次のように 4 問続けて, a1, a2, a3, a4, と解く形態を指す。一方, 交互練習では, ある特定の方略を連続的に使うことは避けて, 異なる方略を交互に使用する。例えば, 方略 a を使う問題に, 方略 b と c を使う問題を混ぜて, a1, b1, a2, c1, と解く形態を指す。

この研究パラダイムでは, しばしば, 学校で行われる算数・数学教育の特徴として, 教育は, ブロック練習中心で進み, 交互練習はまとめ等の機会に僅かに含まれるだけである, という点が指摘される。教科書を調査した Dedrick, Rohrer, & Stershic (2016) や Rohrer, Dedrick, & Stershic (2015) によれば, 教科書の構成は, 「ある 1 つの課題 (例えば, 分数の割り算の手続きの習得) を

* 宮崎大学
〒 889-2192 宮崎県宮崎市学園木花台西 1 丁目 1
t-onoue@cc.miyazaki-u.ac.jp

** 生物科学研究所 (長野県岡谷市)

ターゲットにして、まず、その課題の解決に必要な方略（先の例なら、割る数の逆数をかける）を理解し、次に、その課題の練習問題を数問続けて集中的に解く（つまり、直前に習った特定の方略を連続的に使用するブロック練習）を、課題毎に、順に行う形が基本になっているという。当該の単元の終末には、復習として、その単元で学習した複数の課題の各練習問題を、同一課題の練習問題が続かないようランダムにして解く（つまり、異なる方略を交互に使用する）交互練習形式の練習問題も用意されている。しかし、その復習の中には、その単元で学習した複数の課題の各練習問題を、課題毎に順に解く（例えば、a1, a2, a3, b1, b2, b3, c1, c2, c3のように解く）ブロック練習形式のものが含まれることも珍しくないため、結果的に、交互練習の機会は少ないという。Dedrick et al. (2016) では、中学7年生の数学の教科書6冊を対象に、教科書に掲載されている全練習問題の中で交互練習が占める割合を算出し、平均で11%と報告している。

米国での指摘をもとに、我が国の教科書を概観すると、我が国でも、米国と同様の傾向がみてとれるかもしれない。小学校4年生の教科書を例にとると、小数の単元では、例えば、「小数×1桁の整数」の手続きの習得を課題とし、その方略を理解した後に練習問題を6問続けて解く、或いは、「小数÷1桁の整数」の手続きの習得を課題とし、その方略を理解した後に練習問題を8問続けて解く、というように、課題毎にブロック練習する構成がよく用いられることを確認できる。単元の終末には、復習として、その単元で学習した複数の課題の各練習問題が用意されているものの、その中には、先述した、交互練習形式のものと、ブロック練習形式のものがある。他の単元や、他学年の教科書においても同様の特徴がみてとれることを踏まえると、我が国でも、教育は、ブロック練習中心で進み、交互練習はまとめ等の機会に僅かに含まれるだけであるという点は、ある程度あてはまるかもしれない。

本研究では、近年の米国で発展してきている、「算数・数学科の学習を、ブロック練習と交互練習という観点から捉える」研究パラダイムに依拠して、研究を行う。ブロック練習と交互練習は、一見、練習問題の配列の違いという些細な違いにしか見えないかもしれない。だが、詳細は後述するが、その些細な違いが、学習内容の定着度の明確な差として現れる可能性が示されている。ブロック練習と交互練習という考え方が、我が国では未だほとんど検討されていない現状も踏まえると、この研究パラダイムに依拠して我が国での研

究を展開させることには、これまでに無い視点から教育を再考する契機を作れるという点で意義があると考ええる。よって、以下では、まず、現在までの実証研究を概観し、本研究で検討すべき課題を明確にする。

現在までの実証研究では、第1に、ブロック練習と交互練習の各効果の直接比較が行われ、交互練習がブロック練習よりも学習内容の定着に効果があることが示されてきた。例えば、初期の研究である Rohrer & Taylor (2007) では、大学生を対象に、4つの立体の各体積を求める手続きの習得を課題とし、学習条件として、ブロック練習条件と、交互練習条件を設けている。ブロック練習条件では、4つの立体の各練習問題を、立体毎に、順に解く形で（つまり、ある特定の方略を連続的に使用する形で）学習をさせている。一方、交互練習条件では、4つの立体の各練習問題を、同一の立体の問題が続かないようランダムにして解く形で（つまり、異なる方略を交互に使用する形で）学習させている。学習条件間で学習後のテストの得点に違いがあるかを分析した結果、テストの得点は、交互練習条件の方が、ブロック練習条件よりも有意に高かった。このように学習後のテストにおいて、交互練習条件の得点が、ブロック練習条件の得点よりも高くなるという結果は、小学生や中学生を対象にした後続の研究でも示されている (Ostrow, Heffernan, Heffernan, & Peterson, 2015; Patle, Liu, & Koedinger, 2016; Rohrer, Dedrick, & Burgess, 2014; Rohrer et al., 2015; Taylor & Rohrer, 2010; Ziegler & Stern, 2014)。

第2に、ブロック練習と交互練習を組み合わせる学習した時の効果の検討が、Rau, Alevan, & Rummel (2010) で行なわれている（以後、2つの学習法を組み合わせる条件を複合条件、その条件の効果を複合効果と呼ぶ）。現在の教育を、ブロック練習と交互練習という観点から捉えると、「教育は、ブロック練習と交互練習を組み合わせる」と捉えることができる (Dedrick et al., 2016; Rohrer et al., 2015)。それを踏まえると、ブロック練習と交互練習を組み合わせる学習した時の効果に注目し、その効果を実証的に検討することには、現在の教育にみられる、「ブロック練習と交互練習を組み合わせる」という考え方を、反省的に或いは建設的に検討できるという点で意義があると考えられる。Rau et al. (2010) では、小学5,6年生を対象に、分数の概念を図式化する3つの方法（円グラフ、線分、集合表現）の習得を課題とし、学習条件として、ブロック練習条件と、交互練習条件と、ブロック練習後に交互練習する複合条件を設けている。ブロック練習では、3つの図式化の各練

習問題を、方法毎に、順に解く形で（つまり、ある特定の方略を連続的に使用する形で）学習をさせている。交互練習では、3つの図式化の各練習問題を、同一の方法の問題が続かないようランダムにして解く形で（つまり、異なる方略を交互に使用する形で）学習をさせている。学習条件間で学習後のテストの得点に違いがあるかを分析した結果、テストの得点は、ブロック練習条件と、複合条件の得点が、交互練習条件の得点よりも有意に高かったことが報告されている。現時点では、複合効果の検討は、Rau et al. (2010) に限られている。よって、現時点の実証データは、複合効果については、ブロック練習と交互練習の効果を直接比較した研究の結果を踏まえて、単純に予測できない可能性があることを示している。

ただし、Rau et al. (2010) の実験計画には問題点が指摘されているため、複合効果については、再度検討し直す必要があると考える。具体的には、Shah, Sibbald, Jaffer, Probyn, & Cavalcanti (2016) において、ブロック練習と交互練習の直接比較を行ってきた研究 (e.g., Rohrer & Taylor, 2007) では、学習時に必要とされる次の2つの段階、つまり、(1) 学習する課題についての基本的な知識を理解する段階と、(2) その知識を活用する練習を行う段階 (Norman, 2009) の両方が含まれていたが、Rau et al. (2010) では、前者の (1) の段階が含まれていなかったことが指摘されている。すなわち、Rau et al. (2010) では、学習開始後すぐに (2) の段階を導入したことが（つまり、学習開始後すぐに3つの図式化の各練習問題をブロック練習か交互練習で学習させたことが）、交互練習条件の学習を困難にさせた可能性がある一方で、ある特定の方略を連続的に使用し学習するブロック練習が、(1) の段階を補完する役割を果たし（つまり、学習者の中に、当該の図式化に関する基本的な知識を形成する役割を果たし）、そのことが、ブロック練習が含まれる2条件の成績を高めた可能性が指摘されているのである。学習時には、(1) と (2) の両段階が必要とされること (Norman, 2009) を踏まえると、それら両段階が保証された実験計画の中で、複合効果を検討する必要があると言える。

そこで本研究では、学習時に必要とされる2つの段階（上述した (1) と (2) ; Norman, 2009) が保証された実験計画の中で、ブロック練習と交互練習について、それぞれの単独効果、及び、それらを組み合わせた時の複合効果を比較検討することを目的とする。本研究では、その目的を達成する方法として、Rohrer & Taylor (2007) の実験計画を拡張する形で、検討を行う方法を

採用する。Rohrer & Taylor (2007) の実験計画は、(1) と (2) の両段階が含まれる形で構成されるため (Shah et al., 2016)、彼らの実験計画を基にすれば、Rau et al. (2010) に指摘される問題点を改善する形で複合効果を検討できると考えた。Rohrer & Taylor (2007) では、学習を1週間の間隔を空けて2回行う際、次の2つの条件、つまり、2回ともブロック練習で学習する条件と、2回とも交互練習で学習する条件を設けていた。この2つの条件に、本研究では、複合効果を検討する条件として、1回目をブロック練習で学習し2回目を交互練習で学習する条件と、1回目を交互練習で学習し2回目をブロック練習で学習する条件を加えた。先行研究のRau et al. (2010) では、複合効果を調べる条件として、ブロック練習から交互練習という条件のみを設けていたが、複合効果については、交互練習からブロック練習という条件を設けることもできる。それら2つの条件を設定し比較することは、複合効果について、ブロック練習と交互練習の実施順序という観点から検討できる利点があると考え、本研究では2つの条件を設けた。学習内容の定着度を測るテストは、2回目の学習の1週間後に実施し、4つの条件間でテストの得点に違いがあるかを分析する。

本研究では、上記の目的の検討に加えて、確信度判断の正確性、に着目した検討も行う。確信度判断とは、テスト時に自身が想起した内容（解答）が正しいと思う度合いについての自己判断、として概念化されるものである (金城, 2008; 村山, 2009)。その測定は、自身が想起した内容が正しいと思う度合いを、多段階評定を用いて判断させる方法で行われることが多い。分析では、実験参加者による自己判断が、正確か否か、が検討される。本研究のような条件間比較の実験計画では、条件毎に、確信度判断と実際のテストの得点との関連を、相関分析で検定する方法が用いられる (Perfect, 2002)。

本研究では、学習内容の定着に効果を持つ学習法で学習した条件ほど、学習者が、確信度判断を正確に行える、との立場に立つ。確信度判断の知見の蓄積がある再認記憶の研究では、確信度判断の正確性の高さには、その分野に関する知識の豊富さが影響することが示されている (金城, 2008; 高橋, 1998)。それを踏まえると、定着に効果を持つ学習法で学習した条件ほど、学習者が、テスト時に自身が想起した内容の正しさの度合いを、正確に把握できる可能性を予測できる。なぜなら、定着に効果を持つ学習法で学習した条件ほど、学習者の内に残る知識が豊富になり、自身の想起内容の正しさを判断する手がかりが増えると考えられるか

らである。よって本研究では、学習内容の定着度に関する条件間の差を確認した後に、「定着に効果を持つ学習法で学習した条件ほど、確信度判断と実際のテストの得点の関連が強まる」との仮説が支持されるかを検証する。

学習内容の定着に効果を持つ学習法は、確信度判断の正確性を高める、という視点は、従来の研究には無い視点であり、かつ、検討に値する視点であると考ええる。確信度判断という概念は、これまで主に再記憶の研究（目撃証言といった応用的な領域を含む）で扱われており（村山, 2009）、本研究のような学習法の研究ではほとんど扱われていない。だが、この概念を、学習法の研究の文脈に持ち込むことは、学習者が、学習が終了し一定の時間が経過した時点で、学習した内容を想起した際に、自身の想起内容の正しさについての判断を、正確にできるか否かを問題にすることを意味する。そのようなメタ認知的な判断の正確性は、その後の学習行動を的確に計画できるか否かに直結し得るため、重要視する必要性が指摘されている（Soderstrom & Bjork, 2015）。それを踏まえると、学習法の研究の文脈で、確信度判断の正確性に着目し、その正確性に影響する要因を特定しようとする試みには、意義があると言えよう。

方 法

実験参加者

大学1年生68名が実験に参加した。この68名のうち、66名（男性30名、女性36名、平均年齢18.83歳、 $SD=0.73$ ）が分析対象となった。分析対象から外れた2名のうち、1名は、2回目の学習の直前に、参加を辞退したため、分析対象から外れた。この者は、実験条件1

（Table 1）に割り当てられていた。もう1名は、1回目の学習時に、コンピュータの画面上で学習プログラムが動いている最中に、実験者に質問を行った。その間、学習プログラムは動いており、一部分の内容を学習できなかったため、分析対象から外した。

実験課題

Rohrer & Taylor (2007) と同様に、4つの立体について、体積を求める各手続き（公式）を習得することを課題とした。4つの立体は、[A] ウェッジ、[B] 回転楕円体、[C] 球錐、[D] 半円錐である（Figure 1）。なお、全ての実験参加者に対して、1回目の学習の終了後に、本実験課題を学習したことがあるかを口頭で尋ねたが、学習したことがあると回答した者はいなかった。

実験計画

本実験は、学習条件（Table 1の条件1, 2, 3, 4）を独立変数とする実験参加者間計画であった。実験参加者66名は、条件1に16名、条件2に17名、条件3に16名、条件4に17名が、それぞれ割り当てられた。

各条件への割り当て方法を述べる。本実験は、第一著者が、実験室で、1人ずつ個別に行った。1人あたりの実験時間として、1週間の間隔を空けて計3回、それぞれ約45分ずつ時間を確保する必要があったため、同時期に実験を行える人数に限りがあった。そこで、実験参加者を、幾つかのグループに分け、時期をずら

Table 1 本研究で設定した実験条件

条件	学習1	学習2	テスト
条件1 (16名)	ブロック練習	ブロック練習	テスト
条件2 (17名)	ブロック練習	交互練習	テスト
条件3 (16名)	交互練習	ブロック練習	テスト
条件4 (17名)	交互練習	交互練習	テスト

注) 学習1, 学習2, テストの間は、それぞれ1週間空いている。

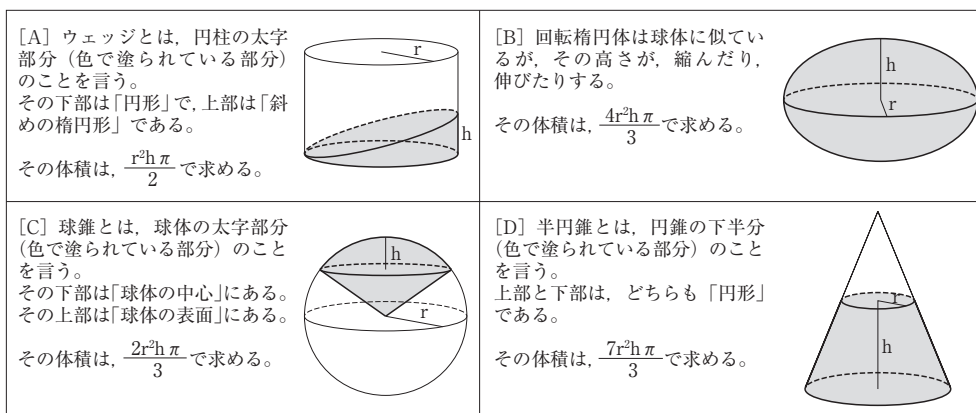


Figure 1 本研究で実験課題として使用した4つの立体

して、グループ毎に順に実験を実施することにした。結果的に、実験は3つの時期(201x年の11月(18名参加)と、12月(27名参加)と、201x+1年の1月(21名参加))に分けて行うことになった。参加者を、より均等に、各条件に割り当てるために、各条件への割り当ては、各時期の参加者毎に次の手続きで行った。まず、実験参加への承諾書の中で5教科(国数社理英)を得意な順に並び替えてもらった結果を使い、数学の順位 averages が均等になるように参加者を2群に分け、一方を1回目の学習をブロック練習で行う群、もう一方を1回目の学習を交互練習で行う群、とした。次に、1回目の学習終了後に、群毎に、1回目の学習の成績が均等になるように2群に分けた。そして、1回目の学習をブロック練習で行った群の片方を条件1に、もう片方を条件2に割り当てた。1回目の学習を交互練習で行った群の片方を条件3に、もう片方を条件4に割り当てた。この手続きで、各時期の参加者は、4つの条件のいずれかに割り当てられた。

実験手続き

本実験は、第一著者が、実験室で、1人ずつ個別に行った。参加者は、学習を1週間の間隔を空けて2回行い、2回目の学習の1週間後にテストを受けた(Table 1)。

学習の手続き ブロック練習と交互練習は、学習時に必要とされる2つの段階、つまり、(1) 学習する課題についての基本的な知識を理解する段階と、(2) その知識を活用する練習を行う段階(Norman, 2009)を合

む形で構成された。(1)の段階では、参加者は、説明資料(Figure 2)を基に、学習する立体についての基本的な知識を理解することを求められた。説明資料には、学習する立体について、その定義と、体積を求める際に使用する公式、及び問題への解答例が記載されていた。(2)の段階では、参加者は、各立体の練習問題を解くことを求められた。その数は、各立体の問題が4問ずつの、計16問であった。なお、この練習問題16問は、学習1と学習2で異なっていた。

ブロック練習と交互練習の手続きの詳細について、まず、共通点を述べる。共通点は、使用される教材(説明資料や練習問題)の内容と量は同じという点にある。どちらの学習でも、「1つの立体についての説明資料」, 「練習問題1問」, 「その練習問題の答え」, が一定の順序で、コンピュータの画面に提示され、参加者はそれをもとに学習を行った(Figure 3)。提示時間は、説明資料が45秒、練習問題が40秒、答えが10秒であった。コンピュータ画面の右上には、残り時間を示すタイマーが表示され(Figure 2の右上を参照)、画面が切り替わる時には音が鳴る仕組みになっていた。参加者は、説明資料が提示されたら、それを読み学習した。練習問題が提示されたら、その解答を、解答用紙に記入した。答えが提示されたら、自分の解答が正しいかを目で見て確認した。なお、練習問題を解く際の解答用紙は、1問につきA4用紙1枚を使用する形で、次の3つの欄、つまり、(1) 当該の問題に使用する公式を記入する欄、(2) その公式に数値を代入し計算する欄、(3)

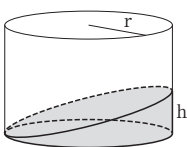
<p>「ウェッジ」の説明</p> <p>「定義と公式」</p> <p>ウェッジとは、円柱の太字部分(色で塗られている部分)のことを言う。</p> <p>その下部は「円形」で、上部は「斜めの楕円形」である。</p> <p>その体積は、$\frac{r^2 h \pi}{2}$で求める。</p>	<p>残り時間 12 秒</p> 
<p>「問題の提示法」</p> <p>半径 (r) = 4, 高さ (h) = 2 の、ウェッジの体積を求めよ。</p> <p>まず、公式を□内に書き、計算し、答えを○内に書け。</p>	<p>「解答の仕方」</p> <div style="border: 1px solid black; padding: 5px;"> <p>公式 $\frac{r^2 h \pi}{2}$ ← (1)^a</p> <p>= $\frac{4^2 \cdot 4 \cdot 2 \pi}{2}$ ← (2)</p> <p>答え 16π ← (3)</p> </div>

Figure 2 学習で使用した説明資料(例として、ウェッジの説明資料)

^a 参加者には、練習問題やテストの問題を解く際には、(1)(2)(3)の順で解答するよう指示した。



Figure 3 ブロック練習と交互練習の手続き

体積を記入する欄 (Figure 2 の「解答の仕方」参照) が設けられていた。参加者は、問題が替わる度に、自分で用紙をめくりながら解答した。答えの提示では、Figure 2 の「解答の仕方」の枠内の情報が提示された。

ブロック練習と交互練習の手続きの差異は、4つの立体の各練習問題を、立体毎に順に解く (つまり、同一の立体の問題を続けて解く) か、同一の立体の問題が続かないようランダムにして解くか、という点にある (Figure 3)。ブロック練習では、1つの立体毎に、説明資料を読

み、練習問題を4問続けて解く形で学習を進めた。交互練習では、まず、4つの立体の説明資料を順に読み、その後は、16問の練習問題をランダムに解く形で学習を進めた。その16問は、次の2つの制約、つまり、⑦各立体の問題1問ずつからなるセットを作りそれを4つ並べる、④セット間の繋ぎ目で同一の立体の問題が続いて出現しない、を満たす形で配置し、順に提示された。

テストの手続き 学習内容の定着度を測るテストで

は、4つの立体の問題を2問ずつ、計8問を解くことを求めた。その8問は、学習時の問題とは異なる新規問題とし、Figure 2の「問題の提示法」の形式で、コンピュータの画面に同時に出题した。8問の並びはランダムとし、画面の左半分と右半分に、それぞれ、縦4問ずつの配置とした。参加者には、どの問題から解いても良いと伝えた。制限時間は8分とし、画面の右上には、残り時間を示すタイマーを表示した。解答用紙として、A3用紙1枚に、8問を解答するための枠をそれぞれ設け、解答を各枠内に記入してもらった。各枠内には、学習時の解答用紙と同様に、3つの欄 (Figure 2の(1)(2)(3))が設けられていた。

確信度判断の測定 テスト後に、各参加者に、自身が想起した解答が正しいと思う度合いを、口頭で尋ねた。7件法が表示された数値軸(「1.全く自信がない」から「7.とても自信がある」)を見せて、自分の認識に近い数字を1つ選んでもらう方法で回答を求めた。

分析方法

テストの得点の分析 本研究の実験課題は、4つの立体の体積を求める各手続き(公式)を習得することであったため、テストの得点は、解答用紙の、「(1)当該の問題に使用する公式を記入する欄 (Figure 2の(1))」に書かれた解答を対象に、算出した。テスト問題8問への各解答が、正しい場合は1点を、正しくない場合は0点を与え、その点数を合算しテストの得点を求めた。なお、当該の公式に含まれる数字とアルファベットが分母と分子の位置に過不足なく表記されていれば、表記の順(例えば、 h と π に関して、どちらが先に書かれてあるか)に関係なく、正答とみなし1点を与えた。テストの得点のレンジは0点—8点である。

テストの得点の分析方法について述べる。本研究では、条件間のテストの得点の差を検討することを主たる目的としているが、分析の際は、学習1と2の各成績(練習問題16問の成績)を含めた経時的変化として分析を行う。学習1と2の各成績は、解答用紙の、「(1)当該の問題に使用する公式を記入する欄 (Figure 2の(1))」に書かれた解答を対象に、テストの得点化の方法と同様の方法で、算出した。得点のレンジは0点—16点である。学習1,2と、テストでは、得点のレンジが異なるため、得点は得点率に変換し、得点率の経時的変化を、線形混合モデルで解析した。

確信度判断の正確性の分析 条件毎に、確信度判断と実際のテストの得点の関連を、Pearsonの相関を用いて検定した。なお、条件1では、1名に確信度判断のデータの欠損がみられたため、その1名は分析から

除外した。

実験参加者を募集する際の手続き、及び倫理的配慮

実験参加者を募集する際の説明は、次のように行った。まず、真の研究目的は伝えずに、本研究は、課題の性質によって学び易さがどう変わるか、を調べる研究であると伝えた。次に、実験に参加した場合には、第一著者が国数社理英の5領域を対象に約10種類の課題を作成しているため、1週間の間隔をあけて3回、実験に参加する度に、ランダムに選んだ1つの課題に取り組んでもらい、その課題の学び易さを評価してもらい、と伝えた。このような説明を行ったのは、参加者に、3回目の研究参加時にテストがあると予想されないようにするためであった。この他、説明時には、実験に参加する場合は、自分が取り組んだ課題の内容を口外しないようお願いした。また、参加者には、お礼として、クオカード1,500円分を渡すことも伝えた。

倫理的配慮として、実験参加者を募集する際の説明時には、研究参加はいつでも辞退できる、データは個人を特定できない形で使用する、等を紙面に明記し、口頭で説明した。実験参加の意思を示した者には、承諾書に署名を求めた。また、実験終了後には、全参加者に、研究結果を説明する機会を設け、その際に、事前の説明時に本研究の真の目的を伝えることができなかった理由も話した。なお、本研究は、第一著者が所属する大学の研究倫理委員会の承認を得て実施された。

結 果

学習中の練習問題の成績と、学習後のテストの成績

条件毎に、3つの時点、つまり、学習1、学習2、テスト、の3時点の得点率を示したのがFigure 4である。得点率の経時的変化を、線形混合モデルで解析した。固定効果として、条件の効果(条件1,2,3,4)と時期の効果(学習1,学習2,テスト)を、変量効果として、参加者内要因(反復測定要因)によるランダム切片効果をモデルに投入した。その結果、交互作用が有意であった($F(6, 186) = 18.68, p < .01$)。そこで、Benjamini-Hochberg法を用いて、単純主効果の多重検定を行った。その結果を下記に記す。

まず、3つの各時点での条件間の差に着目すると、それぞれの時点で条件間の差が確認された。つまり、学習1では、交互練習で学習を行った条件3と4の得点率が、ブロック練習で学習を行った条件1と2の得点率よりも、有意に低かった(いずれも $p < .01$)。学習2では、初めて交互練習で学習を行った条件2の得点率が、他の3条件の得点率よりも、有意に低かった(い

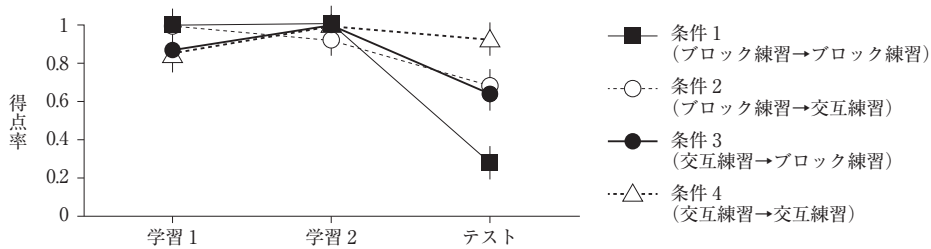


Figure 4 学習とテストにおける各条件の得点率と標準誤差

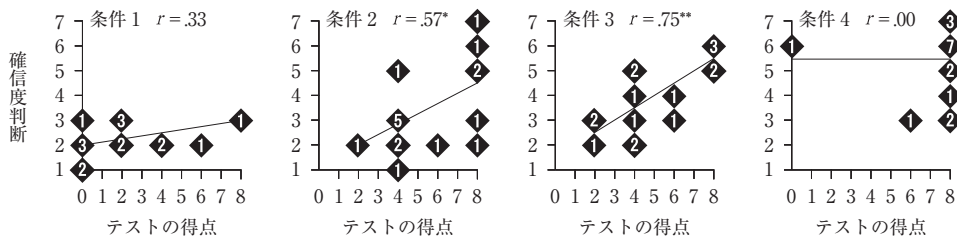


Figure 5 確信度判断と実際のテストの得点の関連

注) 各プロット (◆) の中の数字は、当該のプロットに当てはまる人数である。

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

ずれも $p < .01$ 。一方、学習内容の定着度を測るために実施したテストでは、2回の学習を交互練習で行った条件4と、2回の学習をブロック練習と交互練習を組み合わせて行った条件2及び3と、2回の学習をブロック練習で行った条件1の間にそれぞれ有意差が確認され (いずれも $p < .05$)、得点率は、条件4、条件2及び3、条件1の順で高かった。条件2及び3の間、つまり、2つの学習法の組み合わせの順序を変えた条件間では、有意差は確認されなかった。

次に、3つの時点を通した条件内の差に着目すると、条件1, 2, 3では、テストの得点率が、学習の2時点の各得点率よりも有意に低かったが (いずれも $p < .01$)、条件4では、テストの得点率と学習の2時点の各得点率との間に差はなかった。また、条件4では、学習1から学習2にかけての得点率の増加が有意であり ($p < .05$)、条件3では、学習1から学習2にかけて得点率の増加が有意になる傾向にあった ($p = .052$)。

学習条件毎にみる、確信度判断の正確性

条件毎に、確信度判断と実際のテストの得点との関連を示したのが Figure 5 である。条件毎に、Pearson の相関を用いて検定した結果、ブロック練習と交互練習を組み合わせて学習した2つの条件、つまり、条件2及び3においてのみ、有意な正の相関が認められた。それは、条件2では中程度の正の相関であり ($r = .57$ 、

$p < .05$)、条件3では強い正の相関であった ($r = .75$ 、 $p < .01$)。交互練習単独条件 (条件4) では、 $r = .00$ 、 $p = .99$ 、ブロック練習単独条件 (条件1) では、 $r = .33$ 、 $p = .23$ であり、それぞれ有意な相関は確認されなかった。なお、有意な正の相関が認められた条件2及び3に関して、相関係数の差の検定を行ったが、有意差は示されなかった ($p = .21$)。

考 察

ブロック練習と交互練習の単独効果と複合効果の比較

本研究では、学習時に必要とされる2つの段階、つまり、(1) 学習する課題についての基本的な知識を理解する段階と、(2) その知識を活用する練習を行う段階 (Norman, 2009) が、含まれる実験計画の中で、ブロック練習と交互練習の単独効果と複合効果の比較検討を行った。その結果、複合効果 (条件2及び3の効果) は、ブロック練習の単独効果 (条件1の効果) よりも大きいものの、交互練習の単独効果 (条件4の効果) よりも小さいことが確認された (Figure 4)。特に、本研究では、複合効果を検討する条件として、ブロック練習と交互練習の学習の順序を変えた2条件 (条件2及び3) を設けたが、その条件間のテストの得点に有意差は確認されなかった。それを踏まえると、本研究では、交互練習が学習内容の定着に密接に関係していることを突

き止めた、と言える。なぜなら、交互練習の機会が、条件1で0回、条件2及び3で1回、条件4で2回と、増加するにつれて、テストの得点も増加することが確認されたからである。本研究の意義として、交互練習の機会が増えるにしたがって、学習内容の定着が進むことを示した点を指摘できると考える。

先行研究では、複合効果の検討は、Rau et al. (2010)でのみ行われていた。彼らの結果については、学習に必要なとされる2つの段階のうち(1)の段階、つまり、学習する課題に関する基本的な知識を理解する段階が含まれておらず、そのことが交互練習条件の学習を困難にさせた可能性が指摘されていた(Shah et al., 2016)。彼ら自身も、自身の結果を踏まえて、交互練習は、交互に学習する個々の課題についての知識を得た後に有益な学習法かもしれないと述べている。彼らと本研究の結果を併せて考えると、交互練習単独条件が、複合条件やブロック練習単独条件よりも、学習効果を得るためには、まず、交互練習の中で交互に学習する各課題について、基本的な知識を理解する学習段階を保証し、その上で、その知識を活用する練習を交互練習で行う学習段階に繋げることが、必要だと言えよう。

本研究の結果は、交互練習が学習内容の定着に密接に関係していることを示したが、なぜ、そうなるのだろうか。この点については、理論的には、交互練習は、ブロック練習よりも、想起の努力が求められる分、それぞれの「課題」と「解決方略(或いは課題解決に必要な知識)」との間の強固な記憶が形成されるとの見方がなされている(e.g., Dunlosky et al., 2013)。つまり、交互練習では、複数の課題の各練習問題を、同一課題の練習問題が続かないようランダムにして解く(つまり、異なる方略を交互に使用する)ため、学習者は、問題を見る度に、その問題の解決に必要な方略を、自分の記憶にある複数の方略の中から、逐一思い出さなければならない。この思い出すという認知活動は、想起や検索、或いはテストと呼ばれ(以下、想起と呼ぶ)、その認知活動自体に、思い出そうとした情報についての記憶を強化する役割があることが知られている(e.g., Roediger & Karpicke, 2006)。さらに、想起後にフィードバックがあることで、その想起した内容についての記憶はより一層強化されるという(Storm, Friedman, Murayama, & Bjork, 2014)。この観点から交互練習の学習手続き(Figure 3, 下段)を捉え直すと、問題提示(40秒)が「当該の問題の解決に必要な方略を想起するための時間」に、答え提示(10秒)が「自身の想起の出来を確認するための時間」に位置づいていたとみることができる。ブロック

練習の学習手続き(Figure 3, 上段)においても、交互練習と同様に、問題提示(40秒)と答え提示(10秒)によって学習が進む形になっているが、ブロック練習では、同一課題の練習問題を続けて集中的に解く(つまり、ある特定の方略を連続的に使用する)ため、当該の問題の解決に必要な方略は常に意識下にあり、想起という認知活動に従事する必要性はほとんどなかったかもしれない。条件3に割り当てられたある参加者は、2回目の学習終了後に、実験者に対して、「課題は前回と同じだったけど、今日(ブロック練習)は、前回(交互練習)と頭の使い方が全然違った。単純作業のような感じで、頭をほとんど使わなかった」と自発的に感想を伝えてきた。定着度についての条件間の差(Figure 4)には、2回の学習機会の中で、「それぞれの問題に対して、当該の問題に必要な解決方略を想起し、自身の想起の出来を確認する」という認知活動にどれだけ従事したか、の差が現れているのかもしれない。

本研究の結果は、交互練習が学習内容の定着に密接に関係することを示したが、同時に、交互練習は、それに取り組み始めた初期の段階では、ブロック練習よりも、パフォーマンスが低くなることも示した(Figure 4)。交互練習は、ブロック練習よりも、認知的負荷のかかる学習法であるため(Rohrer, 2009)、学習の初期には、パフォーマンスが低くなり易いかもしれない。だが、学習法の研究では、学習中の出来を意味するパフォーマンスは、長期的な効果を必ずしも反映していないことから、学習中の出来と長期的な効果とを区別する必要性が指摘されている(Soderstrom & Bjork, 2015)。この指摘は、交互練習を実践していく上で、重要かもしれない。それは、本研究や先行研究の結果が示すように、交互練習は、ブロック練習と比べた時、学習後のテストの成績は高くなるものの、学習中の練習問題の成績は低くなる傾向にあるからである。教師や学習者は、学習中のパフォーマンスが芳しくないと感じると、その学習法自体が効果的でないと認識してしまう可能性もある(Kang, 2016b)。よって、交互練習を実践する際は、学習中の出来に過度に囚われないようにすることと、長期的な効果という観点からの評価を行うことの2点が、重要かもしれない。

問題と目的で述べたように、米国では、学校で行われる算数・数学教育に関して、「教育は、ブロック練習と交互練習を組み合わせで展開される」、「ただし、交互練習の機会は少ない」という現状(Dedrick et al., 2016; Rohrer et al., 2015)が指摘されている。本研究の結果は、交互練習の中で要求される認知活動が、学習内容の定

着に大きく貢献することを示した。それを踏まえた時、本研究の結果は、先の現状に対して、ブロック練習と交互練習を組み合わせて学習するという考え方に必ずしも拘る必要はなく、学習者に恩恵をもたらす形で交互練習をどのように導入できるかを積極的に考えることが重要であることを示唆していると言える。ただし、本研究の結果は、本研究の対象者や実験課題に依存している可能性がある。特に本研究では、大学生を対象に実験を行った。ブロック練習と交互練習の直接比較を行った研究では、大学生と小中学生で同様の結果が得られているが、複合効果についても同様の結果が得られるのかを確かめる必要がある。まとめるならば、今後の研究では、第1に、より年少の学習者やスキルが低い学習者を対象者にしたり、或いは、本研究で用いたような手続き的な課題ではなく、概念的な課題やより抽象的な課題を実験課題にして、単独効果と複合効果の検討を行うことが重要である。対象者や、実験課題によって、或いはそれらの要因の組み合わせによって、単独効果と複合効果の現れ方がどのように変わってくるかを検討することで、実際の教育への交互練習の適用範囲や取り入れ方も明確になっていくと思われる。第2に、本研究では、実験条件への参加者の割り当ては、実験課題に関連する教科領域（数学）の得意性についての参加者の主観的評価と、1回目の学習時の成績をもとに行った。この方法は、実験条件の等質性の保証という点では、十分でなかった可能性もあるため、今後の研究ではその点の改良も必要である。第3に、我が国の教育の現状を、ブロック練習と交互練習という観点から正確に捉える実態調査を行う必要がある。我が国の教育の現状を正確に把握できれば、我が国の教育への交互練習の導入方法を的確に検討できることに繋がるだろう。

定着に効果を持つ学習法と確信度判断の正確性の関係

学習内容の定着に効果を持つ学習法で学習した条件ほど、学習者が、確信度判断を正確に行える、との仮説は、条件2及び3と、条件1の比較の範囲においては、支持されたと言える。なぜなら、条件2及び3は、条件1よりも、学習内容の定着度が有意に高く（Figure 4）、かつ、確信度判断とテストの得点の間の有意な正の相関が、条件2及び3には認められたが、条件1には認められなかった（Figure 5）からである。この結果からは、条件2及び3は、条件1よりも、確信度判断を正確に行えている、と言えよう。

しかしながら、学習内容の定着度が最も高かった条件4の結果は、条件2及び3との比較の範囲において、

仮説を支持しなかった。つまり、相関分析の結果は、条件4よりも、条件2及び3の方が、確信度判断の正確性が高いことを示した。ただし、条件4の場合、17名中15名がテストの得点が満点（8点）であり、他の3条件と比べて、満点以外のデータポイントが少ないため、今回のデータでは相関を十分に分析できていない可能性もある¹。その可能性を考慮すると、条件4にみられた結果については、今後、実験参加者を増やす、或いは、満点にデータが集中しないような工夫をする（例えば、学習終了からテストまでの期間を延ばす）、等の改善を施した実験計画のもと再検討した上で、結論する方が良いと思われる。

本研究の結果は、条件2及び3と、条件1の比較の範囲においてのみという限界つきではあるものの、学習内容の定着に効果のある学習法が、副次的に、確信度判断の正確性を高めることを示した。確信度判断の正確性は、その後の学習行動を的確に計画できるか否かに直結する変数になり得ること（Soderstrom & Bjork, 2015）を踏まえると、その正確性に影響する要因を特定した点が、本研究の特色であると言える。今後の研究では、第1に、条件4の結果が仮説を支持しなかったことを踏まえて、交互練習を中心に学習を進めることは、学習内容の定着に効果を持つだけでなく、確信度判断の正確性向上という点でも効果を持つか、について検討する必要がある。第2に、学習法によって、確信度判断の正確性に違いが出る機序の検討が必要である。再認記憶の研究を踏まえると、その背景には、当該の内容に関する知識の豊富さが関係していることが示唆される（金城, 2008; 高橋, 1998）。この点については、例えば、学習者に、学習した内容に関して覚えていることを全て想起してもらう等の方法で知識の豊富さを

¹ 条件4では、テストの得点が0点の者が確信度を6と判断しており、このような顕著な過大判断は4条件の中でこの者にだけみられた。この者は、テスト問題8問全てが誤答と判定されたが、この者の誤答は、1人だけ、他の者の誤答とは質的に異なっていたため、その点を付記しておきたい。今回の学習課題であった4つの立体の体積を求める各手続き（公式）には、共通点部分と差異部分があり、共通点部分は分子の $r^2h\pi$ 、差異部分は分母と分子の数字であった（Figure 1）。この者の場合、8問とも、差異部分の分母と分子の数字は合っていたが、共通点部分の r^2 の²、つまり、2乗が欠如していたため、全ての問題が誤答と判定された。この者のように、差異部分は合っているものの、共通点部分に誤りがあり誤答と判定された者は他におらず、他の者の誤答は、そのほとんどが差異部分に関する誤り（分母と分子の数字の誤り）であり、残りは白紙解答であった。この者の場合、差異部分に関する自身の解答の出来を基準に確信度を判断したが故に、顕著な過大判断となった可能性も考えられる。

測定し、条件間で比較すると良いかもしれない。上記の内容に関する実証的知見を得ることで、学習法と確信度判断の正確性の関係に関する我々の理解が深まり、また、得られた知見は、学習内容の定着と確信度判断の正確性の両側面から効果が見込める教育環境を学習者に提供することに役立てられるだろう。

引用文献

- Booth, J. L., McGinn, K. M., Barbrieri, C., Begolli, K. N., Chang, B., Miller-Cotto, D., ... & Davenport, J. L. (2017). Evidence for cognitive science principles that impact learning in mathematics. In D. C. Geary, D. B. Berch, R. J. Ochsendorf, & K. M. Koepke (Eds.), *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (Vol. 3, pp. 297-325). London: Academic Press.
- Busch, B., & Watson, E. (2019). *The science of learning: 77 studies that every teacher needs to know*. New York: Routledge.
- Dedrick, R. F., Rohrer, D., & Stershic, S. (2016). *Content analysis of practice problems in 7th grade mathematics textbooks: Blocked vs. interleaved practice*. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC.
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*, 4-58. doi:10.1177/1529100612453266
- Kang, S. H. K. (2016a). Spaced repetition promotes efficient and effective learning: Policy implications for instructions. *Behavioral and Brain Sciences, 3*, 12-19. doi:10.1177/2372732215624708
- Kang, S. H. K. (2016b). The benefits of interleaved practice for learning. In J. C. Horvath, J. M. Lodge, & J. A. C. Hattie (Eds.), *From the laboratory to the classroom: Translating science of learning for teachers* (pp. 79-93). New York: Routledge.
- 金城 光 (2008). メタ記憶 太田信夫・多鹿秀継 (編) 記憶の生涯発達心理学 (pp. 232-245) 北大路書房
- 文部科学省 (2017). 小学校学習指導要領 (平成 29 年度告示) 文部科学省 Retrieved from http://www.mext.go.jp/a_menu/shotou/new-cs/1384661.htm (2018 年 1 月 31 日)
- 村山 航 (2009). メタ記憶の測定 清水寛之 (編) メタ記憶—記憶のモニタリングとコントロール (pp. 41-64) 北大路書房
- Norman, G. (2009). Teaching basic science to optimize transfer. *Medical Teacher, 31*, 807-811. doi:10.1080/01421590903049814
- Ostrow, K., Heffernan, N., Heffernan, C., & Peterson, Z. (2015). Blocking vs. interleaving: Examining single-session effects within middle school math homework. In C. Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo (Eds.), *Artificial intelligence in education* (pp. 338-347). New York: Springer. doi:10.1007/978-3-319-19773-9_34
- Patel, R., Liu, R., & Koedinger, K. (2016). When to block versus interleave practice? Evidence against teaching fraction addition before fraction multiplication. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 2069-2074). Austin, TX: Cognitive Science Society.
- Perfect, T. J. (2002). When does eyewitness confidence predict performance? In T. J. Perfect & B. L. Schwartz (Eds.), *Applied metacognition* (pp. 95-120). New York: Cambridge University Press.
- Rau, M. A., Alevan, V., & Rummel, N. (2010). Blocked versus interleaved practice with multiple representations in an intelligent tutoring system for fractions. In V. Alevan, J. Kay, & J. Mostow (Eds.), *Intelligent tutoring systems* (pp. 413-422). Berlin/Heidelberg, Germany: Springer-Verlag.
- Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*, 181-210. doi:10.1111/j.1745-6916.2006.00012.x
- Rohrer, D. (2009). The effects of spacing and mixing practice problems. *Journal for Research in Mathematics Education, 40*, 4-17.
- Rohrer, D., Dedrick, R. F., & Burgess, K. (2014). The benefit of interleaved mathematics practice is not limited to superficially similar kinds of problems. *Psychonomic Bulletin & Review, 21*, 1323-1330. doi:10.3758/s13423-014-0588-3
- Rohrer, D., Dedrick, R. F., & Stershic, S. (2015). Inter-

- leaved practice improves mathematics learning. *Journal of Educational Psychology*, *107*, 900-908. doi:10.1037/edu0000001
- Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science*, *35*, 481-498. doi:10.1007/s11251-007-9015-8
- Shah, R., Sibbald, M., Jaffer, N., Probyn, L., & Cavalcanti, R. (2016). Online self-study of chest X-rays shows no difference between blocked and mixed practice. *Medical Education*, *50*, 540-549. doi:10.1111/medu.12991
- Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science*, *10*, 176-199. doi:10.1177/1745691615569000
- Storm, B. C., Friedman, M. C., Murayama, K., & Bjork, R. A. (2014). On the transfer of prior tests or study events to subsequent study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 115-124. doi:10.1037/a0034252
- 高橋 晃 (1998). 再認の正答率と確信度評定の関連について. *心理学研究*, *69*, 9-14. doi:10.4992/jpsy.69.9
- Taylor, K., & Rohrer, D. (2010). The effects of interleaving practice. *Applied Cognitive Psychology*, *24*, 837-848. doi:10.1002/acp.1598
- Weinstein, Y., Madan, C. R., & Sumeracki, M. A. (2018). Teaching the science of learning. *Cognitive Research: Principles and Implications*, *3*, 1-17. doi:10.1186/s41235-017-0087-y
- Ziegler, E., & Stern, E. (2014). Delayed benefits of learning elementary algebraic transformations through contrasted comparisons. *Learning and Instruction*, *33*, 131-146. doi:10.1016/j.learninstruc.2014.04.006

付 記

本研究は科学研究費助成事業（若手研究（B）, 課題番号：17851628）の助成を受けて実施されました。実験に参加して下さった学生の皆様, ならびに貴重なご指摘を頂いた査読者の先生方に, 御礼申し上げます。

(2019.4.15 受稿, 2020.1.19 受理)

Comparison of Individual and Combined Effects of Blocked and Interleaved Practice: Degree of Retention and Accuracy of Confidence Judgments

TAKAYA ONOUE (UNIVERSITY OF MIYAZAKI) AND YUTAKA IGUCHI (LABORATORY OF BIOLOGY, OKAYA)

JAPANESE JOURNAL OF EDUCATIONAL PSYCHOLOGY, 2020, 68, 122-133

The experiment in the present study compared individual and combined effects of blocked practice and interleaved practice as learning techniques. University students ($N=66$) participated in 2 learning sessions separated by a week, and then, 1 week later, took a memory-recall test. The students were divided into 4 groups: Group 1, blocked practice in both sessions; Group 2, first session, blocked practice; second, interleaved practice; Group 3, first session, interleaved practice; second, blocked practice; and Group 4, interleaved practice in both sessions. Significant between-group differences were found in percentage of correct answers on the test, with Group 4 (interleaved practice in both sessions) being the best performers, followed by Groups 2 and 3, then by Group 1 (blocked practice in both sessions) ($Group\ 4 > 2 = 3 > 1$). These findings suggest that increasing interleaved practice improved retention. Additionally, the participants were asked to judge their confidence in their test performance, using a multistage rating system. When the relation between confidence judgments and test scores per group was analyzed, the results indicated that the participants were more accurate in their assessment of their performance when they used the learning technique that better enhanced retention.

Key Words: learning techniques, blocked practice, interleaved practice, test score, accuracy of confidence judgment