

論文

問題文の長さが誤記憶の修正に及ぼす影響

Length of cue question and memory error correction

高橋 功¹⁾・岩木 信喜²⁾・三浦 春香³⁾・野中 勇樹⁴⁾

Isao Takahashi, Nobuyoshi Iwaki, Haruka Miura, Yuki Nonaka

キーワード：注意資源，誤記憶の修正，過剰修正効果，フィードバック

Keywords：attentional resource, memory error correction, hypercorrection, feedback

要旨：誤った記憶（誤記憶）は，確信度が高いほど修正されやすい（過剰修正効果）。本研究は，誤記憶に対する学習者による注意配分という観点で過剰修正効果を説明できると考え，フィードバックの呈示方法を操作し，手がかりとなる問題文を長文と短文に分けて，誤記憶の修正率の分析を行った。第1テストでは，参加者に一般常識問題に回答させ，問題ごとに確信度を評定させた後，正答のフィードバックを与えた。このとき，継時呈示条件ではパソコンに回答を入力させ，正答のみを視覚呈示した。同時呈示条件では，パソコンに回答を入力させ，参加者の回答と正答とを同時に視覚呈示した。第2テストでは，第1テストと同じ問題に回答させた。24名が参加した実験Ⅰでは，フィードバック方法や確信度に関わらず，長文問題で誤記憶の修正率が有意に高かった。この結果は，長文問題の情報量の多さがもたらした天井効果であると解釈された。そこで12名が参加した実験Ⅱでは，第2テストの実施を予告せず，且つ回答を強制しないことによって，参加者の問題文への注意の低下を狙った。その結果，実験Ⅰとは逆に，短文問題で誤記憶の修正率が有意に高かった。また，長文問題で同時呈示条件の修正率に高い傾向が見られた。これらの結果は，学習者の注意配分が誤記憶の修正に関与することを示唆していると考察された。

問題と目的

例えば，漢字の読み方を覚え違いしているなど，人間の記憶には「誤り」とされる情報（誤記憶）も含まれている。この誤記憶は，日常生活の中で，新しく得た「正しい」とされる情報（正情報）に応じて適宜修正される。直観的に考えれば，誤記憶の存在は，正情報の

1) 山陽学園大学総合人間学部生活心理学科

2) 岩手大学教育学部学校教育科

3) 岩手大学教育学部学校教育科

4) 岩手大学教育学部学校教育科

学習を阻害するものであり、事実の知覚すら歪曲する可能性があることも Loftus (1979 西本武彦訳 1987) の目撃者証言の研究等で知られている。ましてや、誤記憶が強く定着している場合は、正情報の学習はより困難になると考えられる。実際、科学教育の分野では、生徒が日常生活の中で獲得した、科学的には誤りとされる“誤概念 (misconception)”が容易に変容しないことが知られており、それはそうした誤概念が日常経験に起源をもつがゆえに、強い信念となり学習者に定着しているためであると考えられてきた (e.g., Chin & Brewer, 1993; Clement, 1982; Minstrell, 1982; Vosniadou & Brewer, 1992, 高橋, 2000)。

ところが、近年、対連合記憶課題や一般常識問題など比較的単純な記憶学習の研究においては、そのような常識に反する現象が示されてきている。その最初の報告、Butterfield & Metcalfe (2001) では、同じ誤記憶でも、確信度が高い場合、すなわち定着度合いが強いと考えられる場合の方が、正答 (正情報) のフィードバックによって修正されやすいことが示された。この現象を Butterfield & Metcalfe (2001) は、誤記憶の“過剰修正効果 (hypercorrection effect)”と呼んだ。この報告以降、同様の現象が数多くの研究で確認されてきた (e.g., Butler, Fazio, & Marsh, 2011; Butler, Karpicke, & Roediger, 2008; Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009; Metcalfe & Finn, 2011)。

過剰修正効果はどのように説明されるだろうか。これまでは、その説明として、親近性仮説 (Butterfield & Metcalfe, 2001, 2006; Metcalfe & Finn, 2011) と注意捕捉仮説 (Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009) が有力であったと考えられる。親近性仮説では、「確信度の高さはその知識領域の親近性の高さを示しており、そうした領域では正情報によって置き換わる予定の記憶が既に符号化されていることが多いので、修正が容易になる」と説明される。他方、注意捕捉仮説では、「高確信の誤りは一種の驚愕事象であり、注意が自動的に正情報に捕捉されて符号化が促進される」と説明される。

しかし、これらの仮説だけでは説明しきれない現象も確認されている。Butler et al. (2008) の実験では、低確信の誤記憶が比較的高い修正率を示し、高確信の誤記憶が必ずしも高い修正率を示さなかったのである。この研究の実験 I では、40 問から成る一般常識問題 (e.g., 世界で一番長い川は?) が使用された。第 1 テストにおいては、各問に正答 (e.g., ナイル川) とルアー (誤情報) を含む 4 つの選択肢が視覚呈示された。実験参加者は、各問について、正しいと思う選択肢を選んだ後、その確信度を 4 段階で評定した。その後、問題文と正答 (正情報) がフィードバックとして視覚呈示された⁵⁾。この一連の第 1 テスト終了後、5 分間の挿入課題が与えられ、第 2 テストが行われた。第 2 テストは、第 1 テストと同じ問題を手がかり再生テストの形式で回答するものであった。その結果、第 1 テストにおける高確信の誤記憶はもちろん高い修正率 (.93) を示したものの、一方で低確信の誤記憶も比較的高い修正率 (.80) を示した。更に、再テストまで 2 日の期間を設けた実験 2 では、統計的有意差はなかったものの、むしろ低確信の誤記憶の修正率 (.73) の方が、高確信のそれ (.62) より高かった。すなわち、過剰修正効果が認められなかったのである。

この結果について、Butler et al. (2008) は、高確信の誤記憶の修正率が低かった点に注目しており、「保持期間が長くなると、元の誤記憶が出現しやすくなるのではないか」と考察した。後に、Butler et al. (2011) は、120 問から成る、科学的知識を問う手がかり再生

⁵⁾ Butler et al. (2008) では、正答フィードバックのない条件も設定されており、フィードバックの効果と確信度の関連の検討こそが中心的検討課題とされていたが、本研究の議論に直結する問題ではないので、ここでは触れていない。

形式の問題を用いて、第2テストまでの遅延期間を6分にした場合と1週間にした場合とで同じ誤記憶の出現率を比較する実験を行った。その結果、7段階のうち最も確信度が高い場合で、6分後の場合は出現率が.03であったのに対し、1週間後の場合は.23であり、遅延期間が長いと高確信の誤記憶が出現しやすいことが確認された。

他方、本研究では、Butler et al. (2008) の実験における、低確信の誤記憶の修正率の高さの方に注目した。低確信の誤記憶が8割以上修正されたという事実は、親近性仮説や注意捕捉仮説で説明しきれない過剰修正効果のメカニズムが存在することを示唆している。

そしてそれは、学習者の注意資源の配分に関係していると考えられる。というのも、Butler et al. (2008) は、第1テストで参加者に選択肢を視覚呈示しており、この点が他の過剰修正効果に関連する研究と大きく異なるのである。すなわち、その手続きは、学習者が誤記憶に対して十分な注意を向けることを可能にしたのではないかと考えられる。

誤記憶に注意を向けることにはどのような意味があるだろうか。現実生活を顧みれば、間違えやすい敬語や漢字をあえて学習しておくなど、誤記憶に注意を向け、誤情報としてのタグを付与、あるいは誤りであるというメタ認知を高めた上で知識の適正化を図っている例は枚挙に暇がない。つまり誤記憶への注意の配分は「それは誤りである」という認識を高め、記憶の修正を促進するのではないかと考えられる。

ただし、高確信の誤記憶は、そもそも表象上で明瞭に意識されていると考えられる。したがって、誤記憶を視覚呈示されなくても修正されやすいと考えられる。他方、低確信の誤記憶は、表象上で明瞭に意識されていないので、誤記憶の視覚呈示の恩恵を受けることができる。そのように、注意資源の配分という観点から考えれば、過剰修正効果も、低確信の誤記憶で高い修正率が見られたケースもうまく説明することができる。

本研究では、この「注意配分仮説」を検証するため、第1テストにおけるフィードバックの呈示方法を操作した。「口頭報告条件」では、口頭で回答を求めた後、正答のみをモニター上に視覚呈示した。「継時呈示条件」では、モニターを見ながら回答をパソコンに入力させた後、正答のみをモニター上に視覚呈示した。「同時呈示条件」では、モニターを見ながら回答をパソコンに入力させた後、その回答と正答とをモニター上に同時に視覚呈示した。すなわち、口頭報告条件、継時呈示条件、同時呈示条件の順で、誤記憶が明瞭に外在化されており、この順で学習者は誤記憶に注意が向けやすくなる。注意配分仮説が正しければ、口頭報告条件で修正率が最も低く、同時呈示条件で修正率が最も高くなると予測される。また、その条件差は高確信のときよりも低確信のときに大きくなると予測される。

ただし、注意資源の配分は他にも様々な要因の影響を受ける。その一つが問題の性質である。それは学習者の得手不得手や関心によって様々であろう。高いコミットメントがある問題であれば、たとえ低確信の誤答しか思い浮かばなかった場合でも、学習者はそれに注意を向け、明瞭に意識すると考えられる。Iwaki et al. (2013) の実験結果は、これを示唆するものである。この研究の実験2では、330問から成る漢字の音読課題が用いられた。参加者は、第1テストで読み方の口頭再生と確信度の評定を行い、正答フィードバックを受け、挿入課題の後、誤った課題について第2テストを受けた。この第2テストの際に、各課題の実用的価値（「その漢字の読み方が分からないと、将来どの程度困るか」）の5段階評定が求められた。その結果、実用的価値が高いと評定された問題ほど、誤記憶の修正率が高かった。これは、確信度に関わらず、課題の価値の高さから、学習者が誤記憶に

対して注意資源を十分に配分したためではないかと考えられる。

そこで本研究では、手がかりとなる問題文の長さに着目し、問題文の長さを変数とした分析を試みた。通常、長文の問題の読解には、短文の問題よりも、多くの注意資源を要するので、誤記憶への注意が低下すると考えられる。そしてそれは参加者の特性に関わらず比較的一様であると考えられる。したがって、注意配分仮説が正しければ、同時呈示フィードバックの効果は、長文の場合に現れやすく、短文の場合に現れにくいと予測される。すなわち、長文の場合、参加者は問題に注意資源を配分しなければならないので、特に低確信の誤記憶の修正率が低下するだろう。それゆえ、同時呈示フィードバックの効果は低確信の誤記憶の修正において顕著に見られると予測されるのである。他方、読解が容易な短文の場合は、低確信の誤記憶の修正率の低下が長文に比して小さいだろう。それゆえ、同時呈示フィードバックの効果も、長文に比して小さいと予測されるのである。

実験 I

方法

参加者 国立 I 大学の学生 24 名（男性 7 名，女性 17 名，19-23 歳）が実験に参加した。いずれの参加者も、実験内容の十分な説明を受け、それに同意した者であった。実験は、2012 年 6 月～12 月中に行った。フィードバック呈示条件は参加者間変数であり、口頭報告条件が 9 名，継時呈示条件が 8 名，同時呈示条件が 7 名であった。

材料 100 問の一般常識問題（e.g., 問題文「国際連合の本部がある都市はどこ？」，正答「ニューヨーク」）を用いた。この問題群は、川口・清水（1992），ユーキャン就職試験研究会（2012），成美堂出版編集部（2012）から引用し文言を修正したものに、類似の問題を追加して構成された。内容は社会・時事一般，歴史・地理，国語・文学史，芸術・芸能，スポーツ，自然科学など広範な領域に渡るものであった。いずれの問題も，正答は短い単語であった。問題の入れ替えを行ったため，参加者間で問題が若干異なった。

手続き 問題文の呈示順序はパソコンの制御でランダム化し，モニターに問題文と回答欄を 1 試行につき 1 つずつ呈示した。参加者には第 1 テストの後，第 2 テストが実施されることをあらかじめ伝え，なるだけ正答を覚えておくよう教示した。

第 1 テストでは，口頭報告条件においては，問題への回答を口頭で行わせた。継時呈示条件と同時呈示条件においては，キーボードで回答を入力させた後，モニター上からその回答を消去した。このとき，いずれの条件においても，参加者には，必ず何か回答するよう求めた。その後，その回答にどの程度自信があるか，0%（「絶対に間違っていると思う」）から 100%（「完全に自信がある」）で確信度を評価させ，その数値をキーボードで入力させた。確信度を入力させた後，正答のフィードバックを行った。このとき，同時呈示条件では，参加者の回答を 1.5 秒間再提示した後，その下に正答を 4 秒間呈示した。口頭報告条件と継時呈示条件では，正答のみを 4 秒間呈示した。

第 1 テスト終了後，挿入課題として数字のパズルを 5 分間行わせてから，第 2 テストを実施した。第 2 テストは，いずれの条件も同じ手続きであり，第 1 テストと同じ 100 問の問題をランダムな順序で呈示して回答のみを求め，フィードバックは正答のみを 2 秒間呈示した。また，参加者には，思い出せないものに限り，回答しないことを許容した。

結果

一部に不適切な問題があったため、それらは第1テストと第2テストでペアごと除外して分析を行った。その上で、第1テストの誤答を、確信度が0%なら“Guess”，1%~49%なら“Low”，50%~100%なら“High”に分類した。文長については、句読点や括弧は含めず、平仮名、片仮名、漢字、数字を一文字として問題文に含まれた文字数を計数し、その文字数が各問題群内で中央値より大きければ“長文”，中央値未満であれば“短文”に分類した。中央値に該当する問題は分析から除外した。その中央値は20名が26，4名が28であった。これらの分類に基づき、全参加者の誤記憶の総数及び第2テストの正答数（すなわち、修正された誤記憶数），及びその比率（修正率）をTable 1に示した。

なお、第1テストの正答率は、短文の口頭報告条件が $.45 \pm .049$ ($M \pm SEM$)，継時呈示条件が $.42 \pm .038$ ，同時呈示条件が $.57 \pm .052$ ，長文の口頭報告条件が $.47 \pm .044$ ，継時呈示条件が $.34 \pm .044$ ，同時呈示条件が $.47 \pm .061$ であった。

Table 1 条件ごとの第1テストの誤記憶数，第2テストの正答数及び修正率（実験I）

条件		短文				長文			
		Guess	Low	High	合計	Guess	Low	High	合計
口頭報告 <i>n</i> =9	誤記憶数	86	72	66	224	84	88	53	225
	正答数	52	59	59	170	66	77	45	188
	修正率	.60	.82	.89	.76	.79	.88	.85	.84
継時呈示 <i>n</i> =8	誤記憶数	100	74	35	209	124	99	31	254
	正答数	66	56	28	150	99	77	26	202
	修正率	.66	.76	.80	.72	.80	.78	.84	.80
同時呈示 <i>n</i> =7	誤記憶数	62	53	22	137	83	68	23	174
	正答数	51	40	19	110	76	55	19	150
	修正率	.82	.75	.86	.80	.92	.81	.83	.86
合計 <i>n</i> =24	誤記憶数	248	199	123	570	291	255	107	653
	正答数	169	155	106	430	241	209	90	540
	修正率	.68	.85	.88	.77	.85	.85	.86	.84

更に、第2テストにおける誤記憶の修正率を参加者ごとに算出し、その平均値を算出した。また、確信度の高さと修正率の関連を見るために、Goodman-Kruskalの順序連関係数(γ)を参加者ごとに算出し、その平均値を算出した。このとき、参加者によってはデータの分布から γ を算出できないケースがあり、その場合は算出せず、平均値の算出からも除外した。以上の分析結果をTable 2及びFigure 1.に示した。

誤記憶の修正率について、データのなかったセル（すなわち、第1テストで当該確信度の誤記憶がなかったセル）には条件に応じた平均値を代入した上で、逆正弦変換し、フィードバック呈示条件(3)×文長(2)×確信度(3)の三要因の混合分散分析を行った。その結果、文長と確信度の交互作用($F(2, 42) = 4.02, p < .05$)が有意であり、短文における確信度の単純主効果($F(2, 42) = 11.13, p < .05$)，Guessにおける文長の単純主効果($F(2, 42) = 15.37, p < .01$)が有意であった(短文 < 長文)。短文における確信度の単純主効果について多重比較を行ったところ、GuessがLow, Highよりも有意に低かった($p < .00167$, 全体の有意水準を $p < .05$ とした際にBonferroni法により調整された有意水準)。

Table 2 条件ごとの第2テストにおける誤記憶の修正率及び第1テスト時の確信度と誤記憶の修正率の連関係数 (γ)の平均値 (実験 I)

条件	短文				長文				
	Guess	Low	High	γ	Guess	Low	High	γ	
口頭報告	修正率	.62 (.04)	.83 (.06)	.93 (.03)	.69 (.12)	.84 (.05)	.86 (.05)	.93 (.05)	.45 (.21)
	n	6	9	9	9	8	9	9	7
	t				5.75**				2.14+
	g				1.92				.81
継時呈示	修正率	.63 (.07)	.83 (.06)	.83 (.07)	.44 (.17)	.81 (.03)	.83 (.05)	.80 (.07)	.02(.09)
	n	8	8	8	8	8	8	8	8
	t				2.66*				.24
	g				.94				.09
同時呈示	修正率	.81(.09)	.89 (.06)	.86 (.09)	.55 (.31)	.90 (.06)	.85 (.06)	.85 (.07)	.02(.32)
	n	6	7	6	5	6	7	7	6
	t				1.75				.06
	g				.78				.02
合計	修正率	.68 (.04)	.85 (.04)	.88 (.04)	.57 (.10)	.85 (.03)	.85 (.03)	.86 (.04)	.17 (.12)
	n	20	24	23	22	22	24	24	21
	t				5.57**				1.35
	g				1.28				.29

カッコ内は SEM, t は student の t , g は Hedges の g

+ $p < .10$, * $p < .05$, ** $p < .01$

修正率の平均値は参加者ごとの修正率に基づくので Table 1 の修正率と一致しない

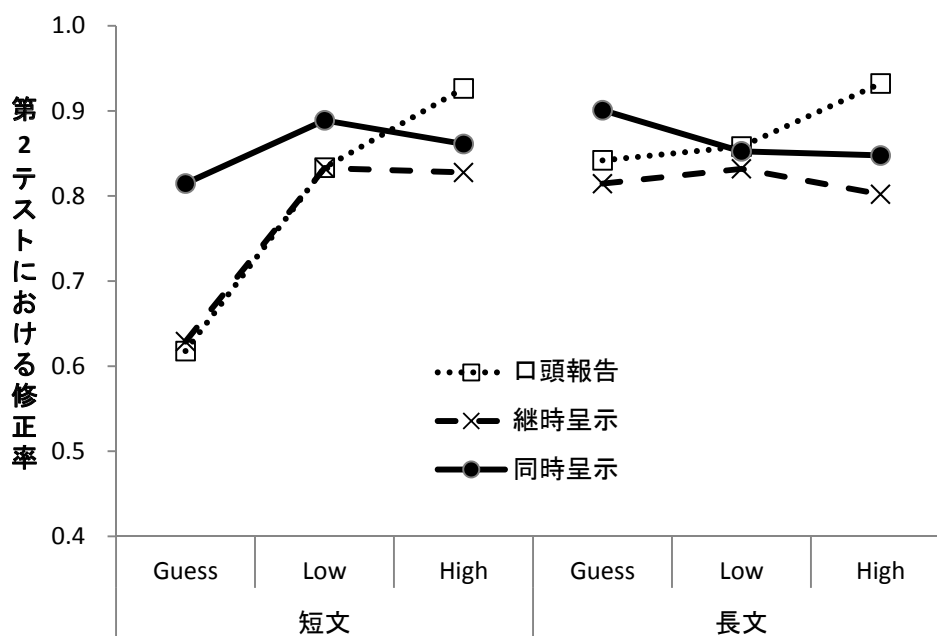


Figure 1. 条件ごとの誤記憶の修正率の平均値 (実験 I)

考察

同時呈示フィードバックの効果が長文の場合に現れやすいという予測に反し、統計的な有意差は検出されなかったものの、むしろ効果が見られたのは短文の場合であった。しかしそれは、長文に効果が現れなかったというよりは、そもそも長文の修正率がいずれの条件においても高かったため、すなわち天井効果となったためであると考えられる。

そしてその長文の優位性は、特に **Guess** 試行において見られた。これもまた、長文の場合に低確信の誤記憶の修正率が低下するという予測とは逆の結果であった。これはなぜだろうか。第1テストの正答率に違いはないので、長文問題の方が内容的に平易であったとは考えにくい。長文を読んだ場合の方が誤答時の驚愕が大きくなるとは考えにくいので、この現象を注意捕捉仮説で説明するのも難しい。確信度の高さを前提とする親近性仮説でも説明できない。

恐らく、含まれる情報量が相対的に多い長文問題は、手がかりの符号化あるいは検索を促進したのではないだろうか。本実験は、他の多くの過剰修正効果に関連する研究と異なり、第2テストの実施をあらかじめ伝えた。それにより参加者は、問題と正答とを連合させることに注意を配分したと考えられる。そうしたとき、情報量の多い長文問題では、特定のキーワードを任意に選びやすい。そして特に **Guess** 試行の場合、参加者は自身の誤記憶にほとんど注意を向けず、問題文に注意を多く配分するので、情報量の多さの優位性が顕著に現れたと考えられる。

また、本実験は、すぐに思いつく回答がない場合でも、必ず何か回答するよう参加者に求めた。これにより、恐らくなかなか回答が思いつかなかったであろう **Guess** 試行では、参加者が問題文をよりよく読み込んだのではないかと考えられる。このこともまた、長文問題における手がかりの符号化の優位性を促進したと考えられる。

実験Ⅱ

実験Ⅰでは、第2テストの実施を予告したことで、参加者に回答を強制させたことが、情報量の多い長文問題の手がかりの符号化を促進し、天井効果の事態を招いた。そこで、実験Ⅱでは、第2テストを予告せず、思いつく回答がない場合は無回答でもよいこととし、再実験を行った。また、剰余変数を排除するため、フィードバック呈示条件を継時呈示条件と同時呈示条件のみとし、参加者内実験とした。更に、継時呈示条件では参加者の回答の代わりにダミーを表示することで、正答フィードバック呈示までの時間統制も行った。

参加者 私立S大学の学生12名(男性4名、女性8名、19-23歳)が実験に参加した。いずれの参加者も、実験内容の十分な説明を受け、それに同意した者であった。実験は、2012年12月～2013年1月中に行った。フィードバック呈示条件は参加者内変数であり、全12名ともに参加した。

材料 110問の一般常識問題を用いた。内容は、実験Ⅰのものをベースとし、同種の問題を追加したものであった。継時呈示条件と同時呈示条件には、パソコンの制御により、55問ずつが参加者ごとにランダムに割り振られた。問題の入れ替えを行ったため、参加者間で問題が若干異なった。

手続き 問題文の呈示順序はパソコンの制御でランダム化し、モニターに問題文と回答欄を1試行につき1つずつ呈示した。参加者には、第2テストが実施されることをそのときまで伝えなかった。

第1テストでは、参加者にキーボードで回答を入力させた後、モニター上からその回答

を消去した。このとき、どうしても思い出せないものに限り、回答しないことを許容した。その後、実験Ⅰと同様にその回答の確信度を評価させ、その数値をキーボードで入力させた。その後、正答のフィードバックを行った。このとき、同時呈示条件では、参加者の回答をモニター上に1秒間再提示した後、その下に正答を3秒間呈示した。継時呈示条件では、参加者の回答の代わりにダミーとして9つのアスタリスク（*****）を1秒間呈示した後、正答のみを3秒間呈示した。

第1テスト終了後は、挿入課題として数字のパズルを5分間行わせてから、第2テストを実施した。第2テストは、いずれの条件も同じ手続きであり、第1テストと同じ110問をランダムな順序で呈示して回答のみを求め、フィードバックは正答のみを1秒間呈示した。また、参加者には、思い出せないものに限り、回答しないことを許容した。

結果

第1テストにおける誤答を、無回答であれば確信度に関わらず“Skip”，確信度が0%なら“Guess”，1%~49%なら“Low”，50%~100%なら“High”に分類した。文長については、実験Ⅰと同様に、中央値に基づいて“長文”と“短文”に分類した。その中央値は9名が26，3名が28であった。これらの分類に基づき、実験Ⅰと同様に、全参加者の誤記憶の総数と修正率をTable 3に示した。

なお、第1テストの平均正答率は、短文の継時呈示条件が $.17 \pm .04$ ($M \pm SEM$)，同時呈示条件が $.20 \pm .03$ ，長文の継時呈示が $.21 \pm .03$ ，同時呈示が $.19 \pm .03$ であった。

Table 3 条件ごとの第1テストの誤記憶数，第2テストの正答数及び修正率（実験Ⅱ）

条件	条件	短文					長文				
		Skip	Guess	Low	High	合計	Skip	Guess	Low	High	合計
継時呈示	誤記憶数	125	27	54	28	234	104	52	64	35	255
	正答数	40	11	30	22	103	41	29	27	18	115
	修正率	.32	.41	.56	.79	.44	.39	.56	.42	.51	.45
同時呈示	誤記憶数	116	41	63	26	246	133	40	60	26	259
	正答数	31	15	33	19	98	37	17	37	16	107
	修正率	.27	.37	.52	.73	.40	.28	.43	.62	.62	.41
合計	誤記憶数	241	68	117	54	480	237	92	124	61	514
	正答数	71	26	63	41	201	78	46	64	34	222
	修正率	.29	.38	.54	.76	.42	.33	.50	.52	.56	.43

$n=12$

更に、第2テストにおける誤記憶の修正率を参加者ごとに算出し、その平均値を算出した。また、確信度の高さと修正率の関連を見るために、Goodman-Kruskalの順序連関係数(γ)を参加者ごとに算出し、その平均値を算出した。このとき、参加者によってはデータの分布から γ 係数を算出できないケースがあり、その場合は算出せず、平均値の算出からも除外した。以上の分析結果をTable 4, Table 5及びFigure 2に示した。

誤記憶の修正率についてデータのなかったセル（すなわち、第1テストで当該確信度の誤記憶がなかったセル）には条件に応じた平均値を代入した上で、逆正弦変換し、文長(2)×フィードバック呈示条件(2)×確信度(4)の三要因の参加者内分散分析を行った。その

結果, 文長と確信度の交互作用 ($F(3, 33) = 5.60, p < .01$) が有意であり, 短文における確信度の単純主効果 ($F(3, 33) = 18.96, p < .01$), High における文長の単純主効果 ($F(1, 11) = 8.01, p < .05$) が有意であった (短文 > 長文)。短文における確信度の単純主効果について多重比較を行ったところ, High が Skip, Guess, Low よりも有意に高かった ($p < .0083$, 全体の有意水準を $p < .05$ とした際に Bonferroni 法により調整された有意水準)。フィードバック呈示条件は, 主効果が有意傾向 ($F(1, 11) = 3.40, p < .10$) を示したのみであった。

また, 参加者の第 2 テストにおける誤答の性質を知るため, 条件ごとに第 2 テストの誤答を分類した。すなわち, 第 1 テストと同じ誤答であった場合は「再誤答」(第 1 テストで無回答だった場合を除く), 各参加者の問題セット内の他の問題の問題文や正答に含まれた単語であった場合は「干渉誤答」, 第 1 テストとは異なる誤答であった場合は「新誤答」(干渉誤答を除く), 何も答えなかった場合は「無回答」として分類を行った。その比率には, 条件間でほとんど違いが見られなかったが, 正答率の低かった長文の継時呈示条件の Low 試行と High 試行では「無回答」の比率が若干高かった。この分類結果を Table 6 に示した。

Table 4 短文問題の条件ごとの第 2 テストにおける誤記憶の修正率及び第 1 テスト時の確信度と誤記憶の修正率の連関係数 (γ) の平均値 (実験 II)

条件		Skip	Guess	Low	High	γ
継時呈示	修正率	.39 (.09)	.40 (.15)	.55 (.11)	.73 (.13)	.29 (.18)
	n	11	7	9	8	12
	t					1.64
	g					.47
同時呈示	修正率	.29(.07)	.36(.14)	.54 (.10)	.82 (.08)	.54 (.07)
	n	11	7	11	9	12
	t					7.29**
	g					2.10

カッコ内は SEM, t は student の t , g は Hedges の g

+ $p < .10$, * $p < .05$, ** $p < .01$

修正率の平均値は参加者ごとの修正率に基づくので Table 3 の修正率と一致しない

Table 5 長文問題の条件ごとの第 2 テストにおける誤記憶の修正率及び第 1 テスト時の確信度と誤記憶の修正率の連関係数 (γ) の平均値 (実験 II)

条件		Skip	Guess	Low	High	γ
継時呈示	修正率	.40 (.06)	.49(.12)	.43 (.08)	.44 (.11)	.05 (.13)
	n	11	8	11	11	12
	t					.38
	g					.11
同時呈示	修正率	.32(.07)	.55(.15)	.65 (.08)	.65 (.13)	.50 (.09)
	n	11	8	9	8	11
	t					5.33**
	g					1.61

カッコ内は SEM, t は student の t , g は Hedges の g

+ $p < .10$, * $p < .05$, ** $p < .01$

修正率の平均値は参加者ごとの修正率に基づくので Table 3 の修正率と一致しない

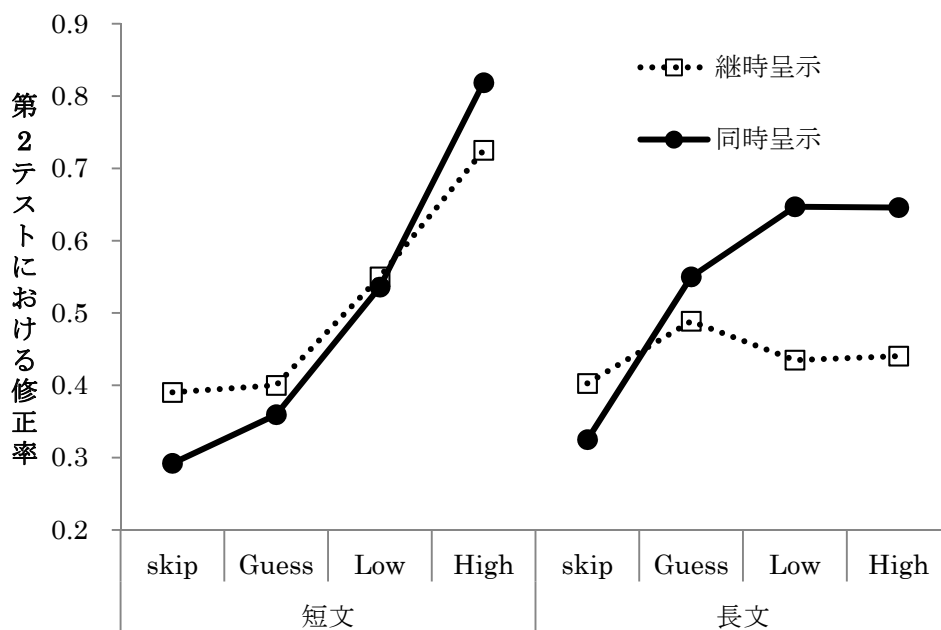


Figure 2. 条件ごとの誤記憶の修正率の平均値 (実験Ⅱ)

Table 6 第2テストの誤答の分類 (実験Ⅱ)

条件	分類	短文				長文			
		Skip	Guess	Low	High	Skip	Guess	Low	High
継時呈示	再誤答	0 (.00)	0 (.00)	3 (.06)	1 (.04)	0 (.00)	3 (.06)	5 (.08)	3 (.09)
	干渉誤答	6 (.05)	3 (.12)	0 (.00)	3 (.11)	3 (.03)	2 (.04)	4 (.07)	1 (.03)
	新誤答	12 (.10)	3 (.12)	6 (.12)	0 (.00)	9 (.09)	2 (.04)	4 (.07)	2 (.06)
	無回答	67 (.54)	10 (.38)	15 (.28)	2 (.08)	51 (.50)	16 (.31)	24 (.38)	11 (.32)
	正答	40 (.32)	11 (.41)	30 (.56)	22 (.79)	41 (.40)	29 (.56)	27 (.43)	18 (.52)
	計	125 (1.0)	27 (1.0)	54 (1.0)	28 (1.0)	104 (1.0)	52 (1.0)	64 (1.0)	35 (1.0)
同時呈示	再誤答	0 (.00)	0 (.00)	3 (.05)	1 (.04)	0 (.00)	1 (.03)	2 (.04)	0 (.00)
	干渉誤答	5 (.05)	1 (.03)	5 (.08)	1 (.04)	5 (.04)	0 (.00)	3 (.05)	3 (.12)
	新誤答	11 (.10)	7 (.18)	6 (.10)	2 (.08)	7 (.06)	12 (.30)	3 (.05)	2 (.08)
	無回答	69 (.60)	18 (.44)	16 (.26)	3 (.12)	84 (.64)	10 (.25)	15 (.25)	5 (.20)
	正答	31 (.27)	15 (.37)	33 (.53)	19 (.74)	37 (.28)	17 (.43)	37 (.62)	16 (.62)
	計	116 (1.0)	41 (1.0)	63 (1.0)	26 (1.0)	133 (1.0)	40 (1.0)	60 (1.0)	26 (1.0)

n=12

カッコ内は条件内での比率

考察

実験Ⅰとは異なり、長文の天井効果は消失した。また、有意傾向に留まるが、フィードバック呈示条件の効果は長文に認められた。ただし、この効果は、どちらかと言えば低確

信ではなく、高確信の誤記憶の方に現れた。また、同時呈示によって修正率が向上したというよりは、長文の継時呈示条件の修正率が低く、過剰修正効果さえも認められなかった。

長文問題の修正率の低さは注意配分仮説による予測の通りであったが、継時呈示条件で過剰修正効果さえも認められなかったことは予測に反する。これはなぜだろうか。第1テストの正答率に違いはないので、長文問題の方が内容的に難解であったとは考えにくい。短文を読んだ場合の方が誤答時の驚愕が大きくなるとは考えにくいので、この現象を注意捕捉仮説で説明するのも難しい。Table 6を見ると、「新誤答」の比率には文長間でも確信度間でも大きな違いはないので、情報量の多い長文問題が参加者の余計な知識を連想させ干渉を起こしたという、親近性仮説のような説明も難しい。また、「干渉誤答」の比率にも大きな違いはないので、情報量の多い長文の問題文自体が干渉を生じさせたとも考えにくい。

サンプルが小さいので慎重に考える必要があるが、長文問題の継時呈示条件では「無回答」による第2テストの誤答が比較的多い。同時呈示条件では文長間にそれほどの違いはないので、参加者が長文の読解への心的努力を特に怠ったとも考えられない。やはり、長文への読解に注意資源が奪われ、自身の誤記憶への注意が抑制されたと考えるのが妥当ではないだろうか。そして、同時呈示フィードバックは誤記憶への注意を高めたと考えるのが妥当ではないだろうか。

総合考察

本研究は、注意資源という観点から、誤記憶を明瞭にフィードバックすることが、学習者の注意を補強し、誤記憶の修正率を高めると予測した。そしてそのフィードバックの恩恵は、誤記憶への注意が低下する低確信の誤記憶、特に長文の問題で顕著化すると予測した。しかし実験Ⅰでは、フィードバックの方法や誤記憶の確信度に関わらず、長文問題の正答率が高かった。この結果は、長文問題の情報量の多さがもたらした天井効果であると解釈された。そこで実験Ⅱでは、参加者に第2テストの実施を予告せず、且つ回答を強制しないことによって、参加者の問題文への注意の低下を狙った。その結果、長文問題の優位性が消失し、予測通り、同時呈示条件において高い修正率が見られた。ただしそれは、低確信の誤記憶ではなく、高確信の誤記憶の修正率を高めたようであった。また、その効果は統計的には有意傾向に留まり、疑問の余地も残すものでもあった。以上の結果は、注意配分仮説を積極的に支持する結果とはいえない。しかし、注意配分仮説による説明と矛盾はしない結果であったし、注意捕捉仮説や親近性仮説では説明し難いものであった。

いずれにせよ、文長が修正率に関わりをもつことは確かであり、この点を示したことに本研究の意義があったと考えられる。もちろん、文長は問題の性質の一側面に過ぎず、必ずしも長文が短文に比べて多くの注意資源を要するとも限らない。漢字の含有率、文節や品詞の数、文章の構造などによっても読解に要する注意の量は異なるであろう。しかし文長は統制が容易な要素であり、注意配分仮説に基づいてフィードバックの効果を検討する際には統制が不可欠であると考えられる。このとき、問題が長文になるほど、更に多くの剰余変数が混入しやすいので、可能な限り、短文にする方向で統制すべきであろう。更に言えば、文章の読解という要素が混入しないような課題、例えば漢字の読み課題などがより望ましい。ただし、その場合は、天井効果を回避するため、回答からフィードバックまでに数唱課題を与えるなど、別の方法で参加者の注意資源を低下させる必要があるだろう。

引用文献

- Butler, A. C., Fazio, L. K., & Marsh, E. J. (2011). The hypercorrection effect persists over a week, but High-confidence errors return. *Psychonomic Bulletin & Review*, **18**, 1238-1244.
- Butler, A. C., Karpicke, J. D., & Roediger, H. L., III. (2008). Correcting a metacognitive error: Feedback increases retention of Low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **34**, 918-928.
- Butterfield, B., & Metcalfe, J. (2001). Errors committed with High confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **27**, 1491-1494.
- Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with High confidence. *Metacognition Learning*, **1**, 69-84.
- Chinn, C. A., & Brewer, F. W. (1993). The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science instruction. *Review of Educational Research*, **63**, 1-49.
- Clement, J. (1982). Student's preconception in introductory mechanics. *American Journal of Physics*, **50**, 66-71.
- Fazio, L.K., & Marsh E.J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review*, **16**, 88-92.
- Iwaki, N., Matsushima, H., & Kodaira, K. (2013). Hypercorrection of high confidence errors in lexical representations. *Perceptual & Motor Skills*, **117**, 1-17.
- 川口 潤・清水博之 (1992) 一般的知識に関する質問項目の難易度および既知感基準表
認知科学会テクニカル・レポート, **21**, 1-16.
- Loftus, E. F. (1979). *Eyewitness Testimony*. Cambridge, MA: Harvard University Press.
(ロフタス, E. F. 西本武彦 (訳) (1987). *目撃者の証言* 誠信書房)
- Metcalfe, J., & Finn, B. (2011). People's hypercorrection of High-confidence errors: Did they know it all along? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **37**, 437-448.
- Minstrell, J. (1982). Explaining the "at rest" condition of an object. *The Physics Teacher*, **20**, 10-14.
- 成美堂出版編集部 (2012). 最新最強の一般常識 一問一答 '14年版 成美堂出版.
- 高橋 功 (2000). 信念と科学的知識の食い違いを子どもはどのように理解しているか: 地球の形の理解を中心にして *発達心理学研究*, **11**, 89-99.
- Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, **24**, 535-583.
- ユーキャン就職試験研究会 (2012). 就職活動の神様の一般常識 これだけ! 一問一答集
自由国民社