

記憶パフォーマンスに及ぼすステレオタイプの 影響に関する実験的検討

高 橋 雅 延

Influence of Stereotype Threat on Optimal Mnemonic Strategies —————

Stereotype threat occurs when individual task performance decreases when a relevant stereotype is salient during the performance. For example, women have more negative beliefs about memory for cars than men, and such gender stereotype might impair performance. Therefore, in this study, we investigated how stereotype threat effects are mediated. It has been suggested that stereotype threat reduces working memory capacity that focuses attention on a goal-relevant task while suppressing task-irrelevant thoughts, which diminishes task performance. To test this theory, we first attempted to directly manipulate the stereotype threat by varying whether the test was characterized as showing gender differences or not. In Experiment 1, female college students ($N=60$, M Age=20.0 years, $SD=1.7$) were randomly assigned to one of two conditions: 30 participants (control group) were informed that past results confirmed the stereotype, and 30 participants (experimental group) were informed that past results did not confirm the stereotype. All participants were required to remember 10 photographs of cars and then conduct the old and new recognition tests. The results indicated that participants in the no-threat experimental group tended to remember more cars than those in the control group. Experiment 2, which followed the same procedure as in Experiment 1, was repeated, except that participants were younger ($n=98$, M age= 9.8 years, $SD=2.5$). Further, older women ($N=44$, M age=74.6 years, $SD=6.6$) were separated into groups of high and low working memory individuals based on their performance on the listening span test. The results of the first experiment were replicated for the older participants but not younger ones. In Experiment 3, the younger participants ($N=269$, M age=18.3 years, $SD=0.5$, also separated into groups of high and low working memory individuals, took the paper-and-pencil version of the mental rotation test in addition to the car recognition task. As a result, participants in the no-threat experimental group tended to show more correct responses in the mental rotation test than those in the control group, but not in the car recognition test. These results suggest that individual differences in working memory capacity did not mediate the observed stereotype threat effect among women. The results are discussed regarding the optimal mnemonic strategy and some methodological problems.

世界経済フォーラム2021年の調査によると、世界156ヶ国のなかで、わが国のジェンダー・ギャップ指数は120位ときわめて低い（内閣府男女共同参画局 <https://www.gender.go.jp/> アクセス年月日2022/09/11）。そのため、わが国の女性は、その活躍が期待されながらも、残念ながら、十分な活躍の場が準備されていない。また、急速に進展する超高齢化社会の中で、「人生100年時代」とも言われ、高齢者の活躍もまた、わが国にとって重要なものとなっている。これらのことから、わが国の女性や高齢者の埋もれた能力を活用し、誰にとっても生きやすい社会の実現のための方策を提唱することが急務となっている。これらに加えて、近年のグローバル化社会において、異質の他者（たとえば、外国人、障がい者など）との共生が求められている中で、女性や高齢者をも含めた共生という社会の新たな価値の実現も目指さねばならない。

これら女性、高齢者、外国人、障がい者などの活躍と共生を阻んでいる大きな要因の一つが、ステレオタイプ（stereotype）と呼ばれる根拠のない固定観念である。ステレオタイプとは、もともと社会心理学の領域で研究されてきた概念であり、特定の集団（たとえば、性別や人種）に対するネガティブな固定観念を指す。たとえば、「女性は男性よりも理数系科目が苦手である」とか、「黒人は白人よりも学業成績が悪い」など、エビデンスのまったく存在しない思い込みのことである。

こういったステレオタイプに関しては、さまざまな研究が行われてきた。その一つの研究の流れは、ステレオタイプがどうやってできあがるのかというステレオタイプの形成過程の研究である（たとえば、McGarty, Yzerbyt, & Spears, 2002）。もう一つの流れは、こういったステレオタイプが、現実の行動において足かせとなり、潜在的な能力を発揮できないというステレオタイプ脅威（stereotype threat）に関する研究である（たとえば、Steele, 2010）。このステレオタイプ脅威という概念を1995年に提唱したのは、自身もアフリカ系アメリカ人であるClaude M. Steeleである。Steele（2010）によれば、我々の社会には種々のステレオタイプが蔓延し

ていて、それが自分に当てはまるかどうかを誰もが敏感に察知しているという。そして、自分に当てはまるステレオタイプと一致した行動を示すかどうかについても、常に他者から観察され、評価されるという脅威に直面しているという。重要な点は、この脅威に気を取られてしまうことで、実際のパフォーマンスが低下し、結果として、ステレオタイプ通りの結果が生じてしまうということである¹。たとえば、「女性は男性よりも理数系科目が苦手である」に関しては、女性が理数系科目を学ぼうとする場合、ステレオタイプ脅威が生じ、そのことに気を取られて、学習に集中できず、結果として、「女性は男性よりも理数系科目が苦手である」というステレオタイプを裏づけてしまうパフォーマンスを示すというのである（当該分野のレビューは、Picho, Rodriguez, & Finnie, 2013などを参照）。

もしそうであるとすれば、このステレオタイプ脅威を取り除けば、そのパフォーマンスはステレオタイプ脅威がある場合に比べて、向上することが予測できる。たとえば、ある研究（Spencer, Steele, & Quinn, 1999, Exp.2）においては、女子学生32名と男子学生24名に対して、数学の問題（高校レベルの代数学、幾何学、三角関数など）に解答させる際に、それぞれのグループに2種類の教示を与えた。すなわち、過去のテスト結果を引き合いに出して、「この数学の問題には明確な性差が見られる」という性差を強調する教示（統制条件）か、「この数学の問題に性差が見られるかどうかは論争中であって明確ではない」という性差なし教示（実験条件）であった（実際には、全員にレベルの等しい2種類のテストを受けさせ、テストごとに、この2種類の教示のいずれかが与えられた。つまり、実験操作としては、参加者内要因であった）。その結果、性差を強調する教示（統制条件）のもとでは、成績において明確な性差が認められた。一方、性差が明確ではないという教示（実験条件）のもとでは、女性の成績が向上し、パフォーマンスの性差が消失した（興味深いことに、男子学生は性差がないという教示のもとでは、それが存在するという教示よりもわずかながら成績が下がった）。この結果から、ステレオタイプ脅威が実際のパフォー

マンスに影響を与えていることが実証されたのである。他にも、数多くの研究が行われる中で、さまざまな種類のステレオタイプ脅威が存在していることが明らかにされている。

本研究で扱うのは、これらのうち、記憶能力に関するステレオタイプ脅威である。すなわち、そのようなステレオタイプには、「高齢になると（若者に比較して）記憶力が悪くなる」という年齢差に関するものもあれば、「女性（男性に比較して）車の記憶力が悪い」という性差に関するものが代表的なものである。本研究の立場は、これら記憶能力に関するステレオタイプの存在のために、高齢者や女性にとっては、それがステレオタイプ脅威となり、本来、個人がそなえている記憶能力を十分に発揮できないというものである。

実際、加齢にともなう記憶力減退というステレオタイプ脅威に関しては、このステレオタイプを除去することで、実際の記憶パフォーマンスが向上することが多くの研究で明らかにされている。たとえば、Rahhal, Hasher, & Colcombe (2001, Exp.2) の研究では、48名の高齢者（年齢のレンジ61-75歳、平均年齢69.38歳）と48名の大学生の若齢者（年齢のレンジ17-24歳、平均年齢19.54歳）を対象として記憶実験を行った。彼らの実験で使われた材料は、たとえば「ダチョウの卵をゆで卵にするには4時間もかかる」といったトリビア（豆知識）に関する文が60文であった。このトリビアには正しいトリビアと誤ったトリビアが含まれ、一つ一つのトリビアの提示後に、その正誤のフィードバックも行われた。実験の最初に、高齢者と若齢者をそれぞれ半数ずつの2群（統制群、実験群）に分け、異なる教示を与えた。統制群である記憶教示群には、記憶実験で与えられる典型的な教示として、「これは新しい事項を覚えることができるかどうかを調べる記憶テスト」であるという教示を与えた。これに対して、実験群であるクイズ教示群には、「提示されるトリビアを学ぶクイズ」であるというように、記憶を強調しない教示を与えた。テストはトリビアそのものを再生するのではなく、先に提示されたトリビアに加え、初めて見るトリ

ピアなどが1つずつ見せられ、それぞれの正誤や「初めて見た」などを判断するという形式で、間接的に記憶を測定した。その結果、通常の記憶教示のもとでは、若齢者の正答率よりも高齢者の正答率が悪かったが、クイズ教示の場合、その差は消失したのである。このことから、記憶の年齢差に関するステレオタイプ脅威が、本来の高齢者の記憶パフォーマンスに影響を与えていることが示されたのである。

この研究では、ステレオタイプ脅威と記憶パフォーマンスとの関連性を明らかにしたとはいえ、ステレオタイプ脅威を直接に操作しているわけではないし、また、記憶パフォーマンスそのものも直接に調べているわけではない。そこで、Hess, Auman, Colcombe, & Rahhal (2003)の研究では、高齢者48名（年齢のレンジ62-84歳、平均年齢70.8歳）と大学生の若齢者48名（年齢のレンジ18-30歳、平均年齢19.3歳）に対して、最新の新聞記事の一節と称したメッセージを読ませることで、ステレオタイプ脅威を直接に操作すると同時に、記憶パフォーマンスも直接に調べた。ステレオタイプ脅威を浮き彫りにさせるネガティブ・メッセージでは、「高齢になるにつれて記憶力が低下するので、家族や友人、それに記憶ツールの助けを借りることが不可欠である」というものであった。これに対して、そのようなステレオタイプ脅威を取り除くポジティブ・メッセージでは、「高齢になるにつれて記憶力が低下するものの、個人の力や環境の力を借りれば、かなりの程度、低下を避けることができる」というものであった。こうして、30語の単語を2分間かけて覚えさせ、直後に自由再生を求めた。その結果、ネガティブ・メッセージを読んだ場合、若齢者よりも高齢者の記憶パフォーマンスの方が悪かったのに対して、ポジティブ・メッセージを読んだ場合、若齢者と高齢者の間には有意差が認められなかった²。

この研究をはじめとして、多くの研究では、若齢者と高齢者を比較している。もちろん、教育歴や健康状態等を両グループで可能な限りそろえているとはいえ、それぞれのグループの人生経験（職業、子育てなど）の違いなどを考えるのならば、単純に若齢者と高齢者を比較することには

問題が残る。そこで、参加者を高齢者だけとしたのが、Hess, Emery, & Queen (2009)の研究である。この研究では、高齢者82名（年齢のレンジは60～86歳，男性39名，女性43名）に対して，2秒間に1語ずつ視覚提示によって50語を覚えてもらった。統制群の教示では，「これは記憶の年齢差を調べる課題で，この課題は若年者よりも高齢者の成績は悪い」というものであった。一方，実験群に対する教示では，記憶ということばを使わずに「これは言語情報の処理に関する個人差を調べる課題で，この課題は年齢にかかわらず，高齢者でもうまくできる」というものであった。こうして，50語を記憶した10分後に，提示した単語と提示していない単語を1語ずつランダムに提示して，「あった」か「なかった」かを判断する再認テストを行った。この際，統制群と実験群をさらに2つの下位グループに分け，答える時間が2500msという制限時間のある場合（ステレオタイプ脅威をより強く感じさせる）と解答時間に制限のない場合を設けた。実験の結果，解答時間に制限のある場合のみ，実験群よりも統制群（つまりステレオタイプ脅威を受けた）の記憶パフォーマンスの方が悪かった。これに対して，解答時間に制限のない場合には，統制群と実験群の記憶パフォーマンスの違いは消失したのである。したがって，高齢者だけを対象にした研究でも，ステレオタイプ脅威が確かに記憶パフォーマンスに影響していることが明らかにされたのである（当該分野のレビューは，Barber, 2020; Lamount, Swift, & Abrams, 2015などを参照）。

しかし，以上述べてきた記憶能力に関するステレオタイプ脅威の研究は，ステレオタイプが喚起されないとステレオタイプ脅威が起こらず，記憶パフォーマンスが悪影響を受けなくなるという現象の記述だけにとどまっているという問題点が指摘できる。ここで明らかにしなければならないのは，ステレオタイプ脅威が起こると，なぜ記憶パフォーマンスが悪くなるのかというメカニズムである。とりわけ，記憶心理学で確立されてきた記憶パフォーマンスに大きな影響を与える効率的な記憶方略(mnemonic strategy)の側面については，何一つ明らかにされてはいない。

目 的

本研究では、ステレオタイプ脅威の状況では、記憶パフォーマンスに関連する効率的な記憶方略がどのような影響を受けるのかという点に注目する。すなわち、記憶能力に関する年齢差や性差といったステレオタイプが喚起され、そのようなステレオタイプ脅威のもとでは、そもそも何が起こり、それが効率的な記憶方略の使用に対して、どのような影響を与えるのかの解明を目的とする。

このような記憶方略の観点からステレオタイプ脅威を検討する際の有力な枠組みが、ワーキングメモリ (working memory) という概念である。ワーキングメモリとは、何らかの課題を遂行する際に、一時的に必要な記憶システム全体を指す (Baddeley, 2007, 2018)。ここではワーキングメモリに関しての詳細を述べることはしないが、ステレオタイプ脅威に関してもっとも重要なのは、ワーキングメモリには容量の限界が存在するという点である。つまり、我々が何らかの課題の遂行時に、その課題に関する記憶や操作などの複数の処理が必要になると、そのすべての処理をこなすことができず、一部の処理は十分に行うことができなくなってしまうという仮定である。

このワーキングメモリ容量とステレオタイプ脅威との関係性を明確にしたのが、Schmaderとその共同研究者たち (Schmader, 2010; Schmader, & Johns, 2003; Schmader, Johns, & Forbes, 2008) である。彼らによれば、ステレオタイプ脅威の状況のもとでは、効率的な記憶方略に必要なワーキングメモリ容量が低下してしまい、その記憶パフォーマンスが悪化するという。ステレオタイプ脅威にさらされると、具体的には、自我を脅かす評価懸念などの課題とは無関連思考の発生とその抑制や、課題実行時のモニタリングが過剰になるために、ワーキングメモリ容量が消費され、効率的な記憶方略の実行が十分にできなくなり、その結果、パフォーマンスが悪

化すると仮定されている。

これらの仮定を踏まえて、本研究では一つの作業仮説を設定する。すなわち、ステレオタイプが喚起されるステレオタイプ脅威のもとでは、自我を脅かす評価懸念が生じ、これが記憶課題とは関連しない無関連思考を誘発するので、それを抑制するためにワーキングメモリが消費される³。このために、効率的な記憶方略に使えるワーキングメモリの容量が不足し、その結果、効率的な記憶方略が使用できず、記憶パフォーマンスが悪くなるという作業仮説である。

この作業仮説の中心となるワーキングメモリ容量を操作・測定する方法は、さまざまなもの（たとえば、二重課題法など）が考えられるが、本研究では、ワーキングメモリ容量の個人差に注目する。先ほどの作業仮説にしたがえば、ステレオタイプ脅威の事態であっても、ワーキングメモリ容量の小さい者よりも大きい者のほうが、ワーキングメモリを活用した効率的な記憶方略を使用できるために、記憶パフォーマンスの低下が消失ないしは減少すると予測することができる。

実験1では、「女性は（男性に比較して）車の記憶力が悪い」というステレオタイプ脅威の現象が存在するかどうかについて、若齢者である大学生を対象に調べる。実験2では、このステレオタイプ脅威と記憶パフォーマンスの関連について、若齢者と高齢者を対象にワーキングメモリ容量の個人差から解明する。実験3では、若齢者だけを対象に、車の記憶課題に加え、女性が苦手とされる空間認知課題であるメンタル・ローテーション課題も加えて、実験2と同様の検討を行う⁴。

実 験 1

目 的

これまで、車の記憶能力は男性に比べて女性の方が悪いという性差の存在することが明らかにされてきた（Davies & Robertson, 1993; Davies,

Kurvink, Mitchell, & Robinson, 1996; McKelvie, Standing, St. Jean, & Law, 1993; Takahashi, Kawaguchi, & Kitagami, 2013)。しかしながら、本研究の立場は、このような「女性は（男性に比較して）車の記憶力が悪い」というステレオタイプの存在が、ステレオタイプ脅威となり、本来の記憶パフォーマンスが阻害されている可能性が高いというものである。

そこで、実験1では、「女性の車の記憶能力は（男性に比べて）悪い」というステレオタイプの存在を前提とした上で、女性の若齢者だけを対象に、ステレオタイプ脅威を除去する教示を与えることによって、ステレオタイプ脅威のある場合よりも、記憶パフォーマンスが良くなるかどうかを調べる。このことによって、「女性の車の記憶能力は（男性に比べて）悪い」というステレオタイプが、もともとの能力差を反映しているのではなく、ステレオタイプ脅威による見かけだけの現象にすぎないことを確認することを目的とした。

方 法

参加者

東京都内のS女子大学の女子学生60名（年齢のレンジは、18～30歳： $M = 20.0$, $SD = 1.7$ ）を30名ずつの2群に分け、統制群と実験群にランダムに割り振った。彼らは全員ボランティアとして小集団実験に参加した。なお、彼らの中に類似の実験を過去に受けた者はいなかった。

実験課題

車の記憶能力の性差について調べたTakahashi et al. (2013) と同じ課題を使用した。

手続き

統制群と実験群に与える教示が異なる以外は、基本的に、Takahashi et al. (2013) と同じ方法であった。すなわち、実験の最初に、統制群には「過去の研究で認められた女性の方が男性よりも車の記憶成績が悪いという性差の確認が研究目的である」と伝えた。一方、実験群には「過去の研究で

は車を真横から映したものが使われていたので女性の方が男性よりも車の記憶成績が悪かったが、今回のように、車を斜め上から映した写真では女性と男性の間には性差が認められないことの確認が研究目的である」と伝えた。

そして、スクリーンに1枚あたり7秒ずつ、合計10枚の写真を見せた。その直後に、再認テストとして、実験中に見せたターゲット写真(10枚)と見せていないディストラクタ写真(10枚)をランダムな順序で1枚ずつ見せて、「あった」か「なかった」かの判断を求める再認とその判断の確信度を5段階評価で求めた。ターゲット写真とディストラクタ写真は各群の半数ごとにカウンターバランスを行った。

なお、実験の最後に、参加者全員に対して、「車の記憶は女性よりも男性の方が優れると言われることに対して、どの程度、同意するか」について5段階評定(1が「まったく同意しない」、5が「まったく同意する」)を求めた。あわせて、実験の実施時点での自動車の運転免許の有無もたずねた。

結 果 と 考 察

結果の分析法

Takahashi et al. (2013)と同様、ヒット率と虚再認率を求め、これらの値から d' 値 (Elliott, 1964) を算出した⁵。以下のすべての統計的分析の α レベルは、特別な言及がない限り、5%とした。

ステレオタイプへの同意の程度と運転免許保有者数

車の記憶の性差に関する同意の程度について、まず、統制群 ($M=3.83$, $SD=0.70$)、実験群 ($M=3.50$, $SD=1.01$) 別に、平均評定値をチャンスレベルと比べたところ、いずれも有意差が認められた(統制群 $t=6.53$, $df=29$, $SE=0.13$: 実験群 $t=2.72$, $df=29$, $SE=0.18$)。また、統制群と実験群の同意の程度の間には有意差は認められなかった ($t=1.49$, $df=58$, $SE=0.22$, $n.s.$)。すなわち、実験参加者が「女性の車の記憶が悪い」というステレオタイプ

を有していることが明らかとなった。ただし、全員が実験終了後にたずねられているために、直前の実験成績から、このような認識を強めていた可能性は否定できないので、結果の解釈には注意が必要である。

また、自動車の運転免許の保有人数は、統制群が7名、実験群が12名であり、両群の間には有意な人数の偏りは認められなかった ($\chi^2=1.93$, $df=1,93$, $n.s.$)。

再認成績

Table 1は、再認テストにおいて、統制群、実験群別に、ヒット率、虚再認率、 d' 値を示したものである。統制群と実験群ごとに、 d' 値をチャンスレベルと比較した結果、統制群 ($t=-10.25$, $df=29$, $SE=0.10$, $p<.05$) も実験群 ($t=-7.78$, $df=29$, $SE=0.10$, $p<.05$) も、いずれもチャンスレベルより有意に低いことが明らかとなった。また、統制群よりも実験群の方が、その成績がよくなる傾向が認められた ($t=1.71$, $df=58$, $SE=0.15$, $p<.10$)。

Table 1 Mean hit, false alarm rates, d' scores in car recognition as a function of group in Experiment 1.

	Hit	False alarm	d'
Control	0.28 (0.12)	0.49 (0.16)	-0.55 (0.56)
Experimental	0.35 (0.12)	0.48 (0.16)	-0.30 (0.56)

Note. Standard deviations are in parenthesis.

これらの結果から、ステレオタイプ脅威を打ち消す教示の効果が認められる傾向が確認された。したがって、車の記憶に関するステレオタイプが、もともとの能力差を反映しているのではなく、ステレオタイプ脅威に基づくという可能性が強く示唆されたと言えよう。

実 験 2

目 的

実験2では、実験1と同じ実験課題と方法を使い、若齢者と高齢者を対象に、ワーキングメモリ容量の測定結果にもとづき、それぞれの年齢群で上位群と下位群に分け、記憶成績の差が認められるかどうかを調べることで、本研究の仮説の検証を行った。

方 法

参加者

東京都内のS女子大学に在籍中の女性若齢者98名（年齢のレンジは、18～43歳： $M=19.8$, $SD=2.5$ ）と、東京都内S区のシルバー人材センターを通じて集められた65歳以上の女性高齢者44名（年齢のレンジは、60～89歳： $M=74.6$, $SD=6.6$ ）であった。若齢者も高齢者もランダムに2群（すなわち、統制群と実験群）に割り振った。その結果、若齢者は統制群49名と実験群49名、高齢者は統制群33名と実験群11名となった⁶。若齢者はボランティアとして研究に参加し、高齢者には3000円の謝礼金を支払った。いずれも小集団実験であり、彼らの中に類似の実験を過去に受けた者はいなかった。

実験課題

＜ワーキングメモリ容量の測定＞

参加者のワーキングメモリ容量を調べるために、遠藤（2013）の日本語版リスニングスパンテストを使った。このテストは、練習が2文あり、その後、2文条件、3文条件、4文条件、5文条件、それぞれが5セットずつあった。ただし、5文条件まで実施すると時間がかかることに加え、参加者の負担が重くなるので、本研究では4文条件までとした。日本語版リスニングスパンテストは、実験者の読み上げる文の文頭の単語を記憶しながら、文の内容が「常識的に合っているかどうか」の正誤判断を行い、所

定の数の文が終わると、すでに聞き終えた記憶すべき単語である文頭の単語を覚えた順番どおりに報告させるものである。たとえば、2文条件では第1文として「ケーキを5つ買い、そのうち2つを食べたら、残りを3人で1つずつ分けることはできない」を聞きながら、文頭の「ケーキ」を覚え、この文の内容の正誤を判断し、続けて第2文として「かかとを上げたとき、地面に接しているのは、5本の指と土踏まずのうち、後者である」を聞きながら、文頭の「かかと」を覚え、この文の内容の正誤を判断し、その直後に、第1文と第2文の文頭の単語であった「ケーキ」「かかと」と思い出さして書かなければならないというものである。

＜無関連思考の抑制能力の測定＞

無関連思考の抑制能力に関しては、日本語版TCAQ (Thought Control Ability Questionnaire) (小林・服部・上野・川口, 2016) を使用した。この質問紙は、オリジナル (Luciano, Algarabel, Tomás, & Martínez, 2005) の25項目のうち、小林ら (2016) が因子負荷量の低い3項目 (「8. 私はいつも自分の考えや行動が適切かどうかを評価している」「14. 私の人生には、『私を困らせる』と感じさせるようなことはほとんどない」「22. 他の人達は自分よりも思考をうまく制御していると思う」) を除外した22項目 (たとえば、「私は心をかき乱す思考を避けられないことがよくある」など) からなっていて、それぞれの項目に対して5段階評定を行い、抑制能力を調べることができるものである。

手続き

実験の最初に日本語版リスニングスパンテストと日本語版TCAQに回答させたあと、実験1とまったく同様に、車の写真の記憶実験を行った。

結 果 と 考 察

結果の分析法

車の写真の記憶に関しては、実験1と同様に、ヒット率とフォールスアラーム率を求め、これらをもとに d' 値を算出した。日本語版リスニング

スパンテストの採点は遠藤（2010）にしたがって行い、ワーキングメモリ（WM）容量の違いによって、リスニングスパン得点が1.5以上を上位群（High WM）、1以下を下位群（Low WM）とした。その結果、若齢者に関しては、上位群（High WM）の統制群26名、実験群15名、下位群（Low WM）の統制群23名、実験群34名となり、高齢者に関しては、上位群（High WM）の統制群8名、実験群6名、下位群（Low WM）の統制群25名、実験群5名となった⁷。日本語版TCAQの採点は小林ら（2016）にしたがった⁸。

再認成績

Table 2は、実験1同様、再認テストにおいて、統制群、実験群別に、ヒット率、虚再認率、 d' 値を示したものである。若齢者の統制群と実験群ごとに、それぞれチャンスレベルと比較した結果、若齢者の統制群（ $t=-12.13$, $df=48$, $SE=0.08$, $p<.05$ ）も実験群（ $t=-21.11$, $df=48$, $SE=0.05$, $p<.05$ ）も、いずれもチャンスレベルより有意に低いことが明らかとなった。また、若齢者の統制群と実験群の間に有意差は認められなかった（ $t=0.79$, $df=96$, $SE=0.09$, $n.s.$ ）。

Table 2 Mean hit, false alarm rates, d' scores in car recognition as a function of group in Experiment 2.

	Hit	False alarm	d'
Young			
Control	0.36 (0.19)	0.51 (0.17)	-0.46 (0.56)
Experimental	0.36 (0.09)	0.56 (0.12)	-0.54 (0.34)
Old			
Control	0.36 (0.18)	0.55 (0.18)	-0.72 (0.84)
Experimental	0.37 (0.45)	0.37 (0.18)	0.02 (0.56)

Note. Standard deviations are in parenthesis.

これに対して、高齢者の統制群と実験群ごとに、それぞれチャンスレベルと比較した結果、若齢者と同様、統制群 ($t=-8.33$, $df=32$, $SE=0.15$, $p<.05$) も実験群 ($t=-2.88$, $df=10$, $SE=0.16$, $p<.05$) のいずれも、チャンスレベルより有意に低いことが明らかとなった。また、高齢者の統制群よりも実験群の方が、その成績がよくなる傾向が認められた ($t=2.69$, $df=42$, $SE=0.27$, $p<.05$)。

高齢者では、統制群と実験群の人数に大きな違いがあるので、解釈には慎重さが要求されるが、この結果からは、ステレオタイプ脅威を打ち消す教示の効果が示唆された。したがって、高齢者においてのみ、ステレオタイプがももとの能力差を反映しているのではなく、ステレオタイプ脅威の効果にすぎない可能性が強いと思われる。

ワーキングメモリ容量と再認成績

Table 3は、若齢者に関して、また、Table 4は高齢者に関して、それぞれの再認テストにおいて、ワーキングメモリ容量の上位群 (High WM) と下位群 (Low WM) 別に、それぞれ統制群、実験群のヒット率、虚再認率、 d' 値を示したものである。

若齢者の d' 値に関して、教示 (統制群、実験群) \times ワーキングメモリ容量 (上位群、下位群) の2要因分散分析を行った。その結果、教示の主効果 ($F(1, 94) = 0.62$, $MSE = 0.21$, $n.s.$, $partial\ eta\ squared = 0.007$)、ワーキングメモリ容量の主効果 ($F(1, 94) = 2.35$, $MSE = 0.21$, $n.s.$, $partial\ eta\ squared = 0.024$)、交互作用 ($F(1, 94) = 0.24$, $MSE = 0.21$, $n.s.$, $partial\ eta\ squared = 0.003$) とともに、すべて有意ではなく、若齢者に関しては、本研究の仮説が支持されなかった。

高齢者の d' 値に関しても、同様の 2×2 の分散分析を行った。その結果、教示の主効果 ($F(1, 40) = 6.93$, $MSE = 0.64$, $n.s.$, $partial\ eta\ squared = 0.148$)、ワーキングメモリ容量の主効果 ($F(1, 40) = 0.14$, $MSE = 0.64$, $n.s.$, $partial\ eta\ squared = 0.003$)、交互作用 ($F(1, 40) = 0.02$, $MSE = 0.64$, $n.s.$, $partial$

Table 3 Mean hit, false alarm rates, d' scores in car recognition as a function of group and working memory (WM) capacity for younger adults in Experiment 2.

	Hit	False alarm	d'
Low WM capacity			
Control	0.35 (0.12)	0.49 (0.16)	-0.36 (0.50)
Experimental	0.35 (0.08)	0.55 (0.12)	-0.51 (0.33)
High WM capacity			
Control	0.37 (0.12)	0.59 (0.13)	-0.56 (0.60)
Experimental	0.36 (0.24)	0.52 (0.17)	-0.61 (0.37)

Note. Standard deviations are in parenthesis.

Table 4 Mean hit, false alarm rates, d' scores in car recognition as a function of group and working memory (WM) capacity for older adults in Experiment 2.

	Hit	False alarm	d'
Low WM capacity			
Control	0.35 (0.18)	0.53 (0.18)	-0.73 (0.90)
Experimental	0.36 (0.15)	0.40 (0.23)	-0.07 (0.64)
High WM capacity			
Control	0.39 (0.19)	0.63 (0.14)	-0.66 (0.64)
Experimental	0.36 (0.16)	0.35 (0.15)	0.09 (0.53)

Note. Standard deviations are in parenthesis.

$\eta^2 = 0.0004$) とともに、すべて有意ではなく、やはり本研究の仮説は支持されなかった。

仮説が支持されなかった理由の一つとしては、小集団実験であったため「自我を脅かす評価懸念」がそれほど強く生じなかったことが考えられる

かもしれない。しかし、実験参加者数が十分に確保できず、さらにはまた、高齢者の中にはリスニングスパンテストの方法の理解が不十分な者も少なからず認められるなど、得られたリスニングスパンテストの得点そのものの信頼性に疑問が残ることも考えられよう。また、車の記憶課題の成績を見ると、実験1も実験2も、チャンスレベル以下という低い成績にとどまっていることも、結果に何らかの影響を及ぼしているのかもしれない。

実 験 3

目 的

実験3では、若齢者だけを対象に、実験2と同様、ワーキングメモリ容量の測定を行うことによって、上位群と下位群に分ける。そして、記憶成績の差が認められるかどうかを調べることで、本研究の仮説の検証を行った。また、車の記憶以外に、性差の認められるとされている空間認知課題（メンタル・ローテーション・テスト）もあわせて実施した。

方 法

参加者

東京都内のS女子大学に在籍中の女性若齢者269名（年齢のレンジは18～20歳： $M=18.3$, $SD=0.5$ ）をランダムに統制群131名、実験群138名に割り振った。彼らは、ボランティアとして研究に参加し、大教室で集団実験を受けた。なお、彼らの中に類似の実験を過去に受けた者はいなかった。

実験課題

実験2と同じ日本語版リスニングスパンテストと日本語版TCAQテストに加え、以下のように、車の再認記憶の材料を変更し、メンタル・ローテーション・テストも実験課題とした。

<再認記憶>

実験1と実験2の車の再認記憶の成績がチャンスレベル以下と非常に悪いことを改善するために、車の写真記憶を難易度の低いものに変更した。

一般に、女性は車の写真の記憶は悪いと言われるのに対して、顔写真の記憶に関しては、男性よりも女性の方がよいことが繰り返し明らかにされている (Lewin & Herlitz, 2002; Lewin, Wolgers, & Herlitz, 2001; McKelvie, 1981; Rehnman & Herlitz, 2007)。実験1と実験2で使用した自動車の写真は斜め上から車全体を撮影したものであった。これに対して、車を真正面から見ると、それが顔に見えることが少なくない (Klatt, Chesham, & Lobmaier, 2016)。そこで、実験3では、車の写真として真正面から撮影した写真を使うことで、全体の成績のレベルを上げることとした。

市販の国産車の写真集より、21枚の車の正面写真を選択し、本実験と同じ母集団で、本実験には参加しない80名の女子学生に、それぞれの車の「記憶しやすさ」と「かわいらしさ」を5段階で評定させた。この評定結果に基づき、「記憶しやすさ」に違いが出ないように、10枚のターゲット写真と10枚のディストラクタ写真を選出し、それぞれの写真をスキャナによってコンピュータに読み込んだ。本来はターゲット写真リストとディストラクタ写真リストのカウンターバランスの必要があるが、今回はカウンターバランスを行わず、固定した。

<空間認知課題>

性差ステレオタイプの1つに、「女性は空間認知能力が（男性に比べて）悪い」という空間認知に関するものがある。実験3では、メンタル・ローテーション・テストを空間認知課題として使うことにした。具体的には、Vandenberg & Kuse (1978) をもとにした Peters (1995) によるペーパー・アンド・ペンシル式のメンタル・ローテーション・テストを使用した。このテストは複数の研究者が使用して一定の妥当性が保障されているものである (椎名・鈴木, 1997, 1998; Peters, Laeng, Latham, Jackson, Zaiyouna, & Richardson, 1995)。

テストは各問4つの立体図形が横並びに印刷され、左側の1つの立体図形と同じものを心的に回転させて4つの中から見つけ出す (チェックを付

ける) という形式である。パート 1 が10問、パート 2 が10問からなり、各パート 3 分間の制限時間が設けられている。ただし、時間の制約と参加者の負担軽減のために、パートを分けずに20問を 3 分間の制限時間で解答させた。採点に関しては、1 問正答ごとに 2 点を与えるので、満点は40点となる。

手続き

リスニングスパンテストの実施方法は実験 2 と同じであった。車の写真の記憶実験では、「車を正面から見ると人の顔のように見えることがある」という説明を行った後、統制群には「車の記憶に関しては、女性の方が男性よりも悪いということが明らかにされている」という教示を与え、実験群には、統制群の教示に加えて「真正面から映した写真に関しては、女性と男性の成績の間に差が認められないことが明らかにされている」という教示を与えた。実験 1 や実験 2 と同様、スクリーンに 7 秒間に 1 枚ずつ見せて、再認テストを行った。

メンタル・ローテーション・テストはパートを分けずに20問が印刷された問題用紙を配布し、3 分間の制限時間内にできるだけ多く回答させた。統制群には、練習後、「このようなテストは女性の方が男性より悪いことが明らかにされている。このことを確認するのが実験の目的である」という教示を与えた。実験群には、練習後、「地図のような平面の場合は女性の方が男性より悪いことが明らかにされている。しかし、このような立体図形に関しては、女性と男性の間に差が認められていないことが明らかにされている。このことを確認するのが実験の目的である」という教示を与えた。

なお、日本語版TCAQは自宅で回答させ、後日、提出させた。

結果と考察

結果の分析法

まず、269名全員のデータをもとに、統制群131名、実験群138名の実験課題ごとの比較を行った。そのあと、リスニングスパンテストを受けた

261名のデータを分析した。この261名のスパン得点は、 $M=2.77$, $SD=0.74$ であった。そこで、平均よりほぼ1SD分高い3.5以上を上位群 (High WM), 2.5以下を下位群 (Low WM) とした。その結果、最終的にデータ欠損のない参加者のうち、上位群 (High WM) の統制群22名、実験群27名、下位群 (Low WM) の統制群42名、実験群46名となった⁹。日本語版 TCAQの採点は、実験2と同様、小林ら (2016) にしたがった¹⁰。車の写真の記憶に関しては実験1や実験2と同様に、ヒット率と虚再認率をもとに、 d' 値を算出した。

再認成績

Table 5の左側3列は、再認テストにおいて、統制群、実験群別に、ヒット率、虚再認率、 d' 値を示したものである。統制群131名と実験群138名ごとに、それぞれチャンスレベルと比較した結果、統制群 ($t=10.80$, $df=130$, $SE=0.12$, $p<.05$) も実験群 ($t=9.63$, $df=137$, $SE=0.12$, $p<.05$) も、いずれもチャンスレベルより有意に高いことが明らかとなった。これは正面写真だと顔と類似しているために、全体の成績レベルが高くなったことが原因だと考えられる。ただし、車の写真の記憶に関して、統制群と実験群の間に有意差は認められなかった ($t=1.30$, $df=267$, $SE=0.17$, $n.s.$)。

メンタル・ローテーション・テストの成績

Table 5の右側の列は、メンタル・ローテーション・テストにおいて、統制群、実験群別に、正答数を示したものである。メンタル・ローテーション・テストの正答数に関して (上記の統制群131名のうち1名のデータ欠損のため)、統制群130名よりも実験群138名の方が成績のよくなる傾向が認められた ($t=1.82$, $df=266$, $SE=0.73$, $p<.10$)。すなわち、実験群の方がステレオタイプ脅威にさらされずに、メンタル・ローテーション・テストの成績が、ステレオタイプ脅威を受けている統制群よりも、高くなるということが示唆された。

Table 5 *Mean hit, false alarm rates, d' scores in car recognition and correct scores in mental rotation as a function of group in Experiment 3.*

	<u>Car Recognition</u>		d'	<u>Mental Rotation</u>
	Hit	False alarm		Correct
Control	0.77 (0.16)	0.23 (0.16)	1.84 (1.42)	11.58 (5.33)
Experimental	0.75 (0.17)	0.25 (0.17)	1.62 (1.36)	12.91 (6.45)

Note. Standard deviations are in parenthesis.

ワーキングメモリ容量と再認成績

Table 6の左側3列は、再認テストにおいて、ワーキングメモリ容量の上位群 (High WM) と下位群 (Low WM) 別に、それぞれ統制群、実験群のヒット率、虚再認率、 d' 値を示したものである。この d' 値に関して、教示 (統制群、実験群) \times ワーキングメモリ容量 (上位群、下位群) の2要因分散分析を行った。

Table 6 *Mean hit, false alarm rates, d' scores in car recognition and correct scores in mental rotation as a function of group and working memory (WM) capacity for younger adults in Experiment 3.*

	<u>Car Recognition</u>			<u>Mental Rotation</u>
	Hit	False alarm	d'	Correct
Low WM capacity				
Control	0.78 (0.16)	0.10 (0.09)	2.37 (0.92)	11.07 (5.91)
Experimental	0.74 (0.18)	0.14 (0.12)	2.06 (0.98)	12.57 (6.02)
High WM capacity				
Control	0.81 (0.16)	0.14 (0.13)	2.37 (1.14)	11.05 (4.40)
Experimental	0.76 (0.18)	0.13 (0.14)	2.27 (1.19)	14.19 (6.24)

Note. Standard deviations are in parenthesis.

その結果、教示の主効果 ($F(1, 133) = 1.78, MSE = 1.10, n.s., partial eta squared = 0.001$), ワーキングメモリ容量の主効果 ($F(1, 133) = 0.36, MSE = 1.10, n.s., partial eta squared = 0.003$), 交互作用 ($F(1, 133) = 0.30, MSE = 1.10, n.s., partial eta squared = 0.002$) は、いずれも有意ではなく、車の記憶に関しては、本研究の仮説は支持されなかった。

ワーキングメモリ容量とメンタル・ローテーション・テストの成績

Table 6の右側の列は、メンタル・ローテーション・テストにおいて、ワーキングメモリ容量の上位群 (High WM) と下位群 (Low WM) 別に、それぞれ統制群、実験群の正答数を示したものである。この正答数に関して、教示 (統制群、実験群) \times ワーキングメモリ容量 (上位群、下位群) の2要因分散分析を行った。

その結果、教示の主効果で有意差が認められた ($F(1, 133) = 4.37, MS = 33.68, p < .05, partial eta squared = 0.032$)。ワーキングメモリ容量の主効果 ($F(1, 133) = 0.80, MSE = 33.68, n.s., partial eta squared = 0.006$), 交互作用 ($F(1, 133) = 0.63, MSE = 33.68, n.s., partial eta squared = 0.004$) は、いずれも有意ではなかった。したがって、メンタル・ローテーション・テストのような空間認知課題の性差もステレオタイプ脅威である可能性が強く示唆された。

これらのことから、車の再認記憶に関しては、全体的に成績のレベルが向上したことの影響なのか、ステレオタイプ脅威の存在を示唆する結果が得られなかった。一方、空間認知課題に関しては、ステレオタイプ脅威の存在を示唆する結果が得られた。一方、ワーキングメモリ容量の個人差に基づく本研究の仮説は明確には支持されなかった。実験2でも指摘したように、「自我を脅かす評価懸念」があまり強く生じなかったことが原因かもしれない。したがって、今後は、たとえば衆人環視のもとで、一人ずつ実験を行うなどして、「自我を脅かす評価懸念」を強める操作が必要であると考えられよう。

全 般 的 考 察

本研究の仮説は、「ステレオタイプが喚起されるステレオタイプ脅威のもとでは、自我を脅かす評価懸念が生じ、これが記憶課題とは無関連な思考を誘発し、記憶方略の使用時に必要となるワーキングメモリの負荷となり、その結果、効率的な記憶方略が使用できず、記憶パフォーマンスが悪くなる」というものであった。しかし、実験2でも実験3でも、この仮説を支持する結果は得られなかった。今後は、この仮説を構成する4つの下位過程（「自我を脅かす評価懸念」「記憶課題とは無関連な思考」「ワーキングメモリの負荷」「効率的な記憶方略」）の一つずつを丁寧に分けて、実験的に検討していくことが重要であると思われる。

一方、教示によるステレオタイプ脅威の操作の結果、実験1では車の記憶、実験2では高齢者の車の記憶、実験3ではメンタル・ローテーション・テストのそれぞれにおいて、ステレオタイプ脅威の存在が示唆された。ステレオタイプ脅威という現象は、状況の認知がさまざまなパフォーマンスに影響を与えるというものである。

記憶心理学者の間では、記憶パフォーマンスを左右するのが記憶方略であるという認識が広く共有されてきた。そのため、これまでの記憶研究では、さまざまな種類の記憶方略（たとえば、リハーサルのタイプや精緻化など）を設定し、どれが記憶パフォーマンスに対して有効であるのかという解明に重点が置かれ、社会心理学の領域で扱われてきたステレオタイプといった社会的要因の検討は行われてこなかった。しばしば引用されるように、1920年代、アメリカのホーソン工場では、さまざまな作業状況が設定され、どれが作業パフォーマンスの上昇に有効であるかが体系的に検討された。その結果、作業状況の違いよりも重要なのが社会的要因（自己注目、人間関係など）であることが明らかにされた（大橋・竹林, 2008; 佐藤, 2007; 進藤, 1978; なお、経営史学会, 2013も参照）。このようなホーソン研

究の結果は、多くの学問の方向性を転換させ、それまで以上に、社会的要因が重視されるようになった。

本研究は、ステレオタイプ脅威という社会的要因を記憶理論に取りこもうという試みである¹¹。今後、本研究のようなアプローチをとる研究が増えることによって、社会的要因を軽視してきた従来の「記憶方略信仰」とでも言うべき偏った記憶理論の問題点が浮き彫りにされ、今後さらに研究を深めることで、既存の知見の再解釈をともなった記憶理論の抜本的な再構築につながることを期待できよう。

さらにまた、社会的要因を重視しているものの、現象の記述にとどまりがちな隣接する心理学の諸領域（社会心理学、教育心理学など）や、他の学問領域（老年学、ジェンダー学、教育学など）に対して、本研究の作業仮説のような記憶方略の視点が加わることで、かつての認知科学がそうであったように、人間理解を共通目標に諸学問が結合した新しい人間理解の総合科学が生みだされる潜在性も秘めていると言えよう。

付 記

本研究は日本私立学校振興・共済事業団の平成30年度（第43回）学術研究振興資金の補助と、平成30年度～令和4年度挑戦的研究（萌芽）（課題番号18K18700）の補助を受けて実施された。実験2が終了した時点では、参加者の人数を大幅に増やし、男子学生や男性高齢者なども含んだ大規模な実験を計画していたが、2020年の新型コロナウイルス感染症の流行により、予定していた実験をすべて断念せざるを得なかった。

本研究の実施に際しては、直接の研究分担者として、心より敬愛する清水寛之先生（神戸学院大学・心理学部）と齊藤智先生（京都大学大学院・教育学研究科）に加わっていただくことができた。また、連携研究者になっていただいた佐藤眞一先生（大阪大学大学院・人間科学研究科）、唐沢穰先生（名古屋大学大学院・環境学研究科）、平井美佳先生（横浜国立大学・

国際総合科学群)には、数々の有益な助言をいただくことができた。私の恩師である梅本堯夫先生は、「日本人が研究を茶道や華道のように単独で実施するものだという思い違いをしている」と、たびたび嘆いておられた。今回、ここに名前を挙げさせていただいた先生たちのおかげで、非常に不十分ながらも、何とか研究を終えることができた。

また、研究成果の貴重な発表の場として、『聖心女子大学論叢』には、これまで29年間の在職期間中に、すでに26本を掲載していただいた。最期となる27本目の掲載にあたり、現在の編集委員会および歴代の編集委員の先生方や、関わっていただいた事務の方々に厚く感謝したい。

世界中のさまざまな著名人の臨終のようすを集めた『人間臨終図鑑』(徳間書店)を書いた山田風太郎が晩年に、心に残っている最期のことばとして、近松門左衛門と勝海舟をあげている。近松門左衛門は臨終の床で、「隠に以て隠にあらず、賢に以て賢にならず、物知りに以て何も知らず、世のまがいもの、からの大和の教ある道々、技能、雑芸、滑稽の類まで、しらぬ事なげに、口にまかせ、筆に走らせ、一生を囀りちらし、今わの際に言うべく思うべく真の一大事は一字半言もなき倒惑」と自分の人生を嘆いたという。大学院の入学から研究生活が始まるとすると、すでに42年もの年月がむなしく流れ去った。この近松門左衛門の臨終のことばに、何一つ成果をあげられなかった研究者のはしくれとして、強い共感の念をいだく。ともあれ、その潔さに感嘆せざるを得ない勝海舟の臨終のことばで終えることにしたい。「コレデオシマイ」。

引用文献

- Baddeley, A. (2007). *Working memory, thought, and action*. Oxford University Press. アラン・バドリー (著) 井関龍太・齊藤智・川崎恵里子 (訳) (2012). ワーキングメモリー思考と行為の心理学的基盤 誠信書房

- Baddeley, A. (2018). *Exploring working memory: Selected works of Alan Baddeley*. Taylor & Francis. アラン・バドリー (著) 佐伯恵里奈・齊藤智 (監訳) (2020). ワーキングメモリの探求—アラン・バドリー主要論文集 北大路書房
- Barber, S. J. (2020). The applied implications of age-based stereotype threat for older adults. *Journal of Applied Research in Memory and Cognition*, **9**, 274-285.
- Davis, G. M., & Robertson, N. (1993). Recognition memory for automobiles: A developmental study. *Bulletin of the Psychonomic Society*, **31**, 103-106.
- Davies, G. M., Kurvink, A., Mitchell, R., & Robinson, N. (1996). Memory for cars and their drivers: A test of the interest hypothesis. In D. J. Herrmann, C. McEvoy, C. Hertzog, P. Hertel, & M. K. Johnson (Eds.), *Basic and applied memory research: Practical applications, vol. 2* (pp. 37-49). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Elliott, P. B. (1964). Table of d'. In J. A. Swets (Ed.), *Signal detection and recognition by human observers* (pp.651-684). New York: Wiley.
- 遠藤香織 (2013). 言語性ワーキングメモリ課題遂行における個人差に関する実験的研究 大阪大学博士論文
- Hess, T. M., Emery, L., & Queen, T. L. (2009). Task demands moderate stereotype threat effects on memory performance. *The Journal of Gerontology: Series B: Psychological Sciences and Social Sciences*, **64B**, 482-486.
- Hess, T. M., Auman, C., Colcombe, S. J., & Rahhal, T. A. (2003). The impact of stereotype threat on age differences in memory performance. *Journal of Gerontology: Psychological Sciences & Social Sciences*, **58B**, 3-11.
- 経営史学会 (監修) 吉原正彦 (編著) (2013). メイヨー＝レスリスバーガー

－人間関係論－（経営学史叢書） 文眞堂

- Klatt, W. K., Chesham, A., & Lobmaier, J. S. (2016). Putting up a big front: Car design and size affect road-crossing behaviour. *PLoS ONE*, **11** (7), e0159455.
- 小林正法・服部陽介・上野泰治・川口潤 (2016). 日本語版Thought Control Ability Questionnaireの作成及び信頼性・妥当性の検討 心理学研究, **87**, 405-414.
- Lamont, R. A., Swift, H. J., & Abrams, D. (2015). A review and meta-analysis of age-based stereotype threat: Negative stereotypes, not facts, do the damage. *Psychology and Aging*, **30**, 180-193.
- Lewin, C., & Herlitz, A. (2002). Sex differences in face recognition – Women's faces make the difference. *Brain and Cognition*, **50**, 121-128.
- Lewin, C., Wolgers, G., & Herlitz, A. (2001). Sex differences favoring women in verbal but not in visiospatial episodic memory. *Neuropsychology*, **15**, 165-173.
- Luciano, J. V., Algarabel, S., Tomás, J. M., & Martínez, J. L. (2005). Development and validation of the thought control ability questionnaire. *Personality and Individual Differences*, **38**, 997-1008.
- McGarty, C., Yzerbyt, V. Y., & Spears, R. (2002). *Stereotypes as explanations*. Cambridge University Press. クレイグ・マクガーディ／ビンセント・Y・ゼルビット／ラッセル・スピアーズ（編著）国広陽子（監修）有馬明恵・山下玲子（訳）（2007）. ステレオタイプとは何か－「固定観念」から「世界を理解する“説得力”」へ－ 明石書店
- McKelvie, S. J. (1981). Sex differences in memory for faces. *Journal of Psychology*, **107**, 109-125.
- McKelvie, S. J., Standing, L., St. Jean, D., & Law, J. (1993). Gender differences in recognition memory for faces and cars: Evidence for the interest hypothesis. *Bulletin of the Psychonomic Society*, **31**, 447-448.

- Merton, R. K. (1949). *Social theory and social structure: Toward the codification of theory and research*. Free Press of Glencoe. ロバート・K・マートン (著) 森東吾・金沢実・森好夫・中島竜太郎 (訳) (1961). 社会理論と社会構造 みすず書房
- 大橋昭一・竹林浩志 (2008). ホーソン実験の研究—人間尊重的研究の源流を探る 同文館出版
- Picho, K., Rodriguez, A., & Finnie, L. (2013). Exploring the moderating role of context on the mathematics performance of females under stereotype threat: A meta-analysis. *Journal of Social Psychology*, **153**, 299-333.
- Peters, M. (1995). *Revised Vandenberg and Kuse Mental Rotations Tests: Forms MRT-A to MRT-D* (Tech. Rep.). Guelph, ON: University of Guelph, Department of Psychology.
- Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R., & Richardson, C. (1995). A redrawn Vandenberg and Kuse Mental Rotations Test: Different versions and factors that affect performance. *Brain and Cognition*, **28**, 39-58.
- Rahhal, T. A., Hasher, L., & Colcombe, S. J. (2001). Instructional manipulations and age differences in memory: Now you see them, now you don't. *Psychology and Aging*, **16**, 697-706.
- Rehman, J., & Herlitz, A. (2007). Women remember more faces than men do. *Acta Psychologica*, **124**, 344-355.
- 作田啓一・井上俊 (編) (1986). 命題コレクション 社会学 筑摩書房
- 佐藤郁哉 (2007). ホーソン実験から組織エスノグラフィーへ—組織研究における2つの起源神話— 一橋商学論叢, **2**, 12-40.
- Scmader, T. (2010). Stereotype threat deconstructed. *Current Directions in Psychological Science*, **19**, 14-18.
- Scmader, T., & Johns, M. (2003). Converging evidence that stereotype

- threat reduces working memory capacity. *Journal of Personality and Social Psychology*, **85**, 440-452.
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, **115**, 336-356.
- 椎名久美子・鈴木賢次郎 (1997). メンタル・ローテーション・テストの問題解決過程に関する考察－ペーパーテストの誤答分析を通して－ 図学研究, **31**, 3-10.
- 椎名久美子・鈴木賢次郎 (1998). 改変版メンタル・ローテーション・テストの設計と誤答分析 図学研究, **32**, 133-138.
- 新藤勝美 (1978). ホーソン・リサーチと人間関係論 産業能率短期大学出版部
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, **35**, 4-28.
- Steele, C. M. (2010). *Whistling Vivaldi: how stereotype affect us and what we can*. W. W. Norton. クロード・スティー爾 (著) 藤原朝子 (訳) (2020). ステレオタイプの科学——「社会の刷り込み」は成果にどう影響し、わたしたちは何ができるのか 英治出版
- Takahashi, M., Kawaguchi, A., & Kitagami, S. (2013). Recognition memory for cars and identification of location: Implications for the unconscious transference of cars. *Journal of Human Environmental Studies*, **11**, 35-41.
- 梅本堯夫 (1987). 認知とパフォーマンス 東京大学出版会
- Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, A group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, **47**, 599-604.

脚注

1. ここで扱っているステレオタイプ脅威は、社会学の分野でしばしば引き合いに出される「予言の自己成就 (self-fulfilling prophecy)」(Merton, 1949) という現象と酷似している。予言の自己成就とは、もともと集団の行動に関する概念であり、「ある状況が起こりそうだと考えて、人びとが行為すると、そう思わなければ起こらなかったはずの状況が、実際に実現してしまう」(作田・井上, 1986, p.80) というものである。たとえば、ある銀行の支払いが不能になるという根も葉もない噂が流れると、預金者が預金を引き出しに殺到するために、結果的に、その銀行が支払い不能に陥ってしまうというような例である。Mertonはこの現象が個人にも生じるとして、試験に強い不安をもっている学生の例をあげている。すなわち、「きっと失敗するにきまっていると思い込んでしまうと、不安な受験生は勉強するよりも、くよくよして多くの時間を浪費し、いざ試験にのぞんでまずいことになる。最初の誤った不安は、いかにももったもた不安に変形してしまう」(Merton, 1949 森ら訳1961, p.384) のである。
2. Hess et al. (2003) の研究では、ネガティブ・メッセージ、ポジティブ・メッセージのいずれのメッセージにも触れない統制群も設けられていた。この統制群でも、若齢者と高齢者の間には有意差が認められなかった。
3. 厳密には課題実行時のモニタリングなども含むべきだが、本研究では記述を簡潔にするために、無関連思考という用語に代表させることにした。
4. 以下の3つの実験研究の倫理的配慮に関しては、日本心理学会の倫理規定および聖心女子大学研究倫理指針、聖心女子大学における「人を対象とする研究」ガイドラインに則り、聖心女子大学研究倫理審査委員会の承認を受けた上で、研究に着手した。収集されたデータはすべて個人の特定を不可能にした上で、電子データとして保存した。

5. ヒット項目と虚再認項目それぞれの確信度も求めたが、実験参加者ごとにそれぞれの項目数のアンバランスが大きいので、信頼性が不十分であると判断し、3つの実験において、いっさいの分析は行わなかった。
6. 本来、高齢者も実験群と統制群に22名ずつに割り振る予定であったが、高齢者の実験補助にあたった研究協力者の手続きミスのため、高齢者の統制群が33名、実験群が11名となった。
7. 若齢者のHigh WM・統制群26名のスパンは $M=2.21$, $SD=0.68$, High WM・実験群15名のスパンは $M=2.40$, $SD=0.60$ であった。また、Low WM・統制群23名のスパンは $M=1.00$, $SD=0.00$, Low WM・実験群34名のスパンは $M=1.00$, $SD=0.00$ であった。一方、高齢者のHigh WM・統制群8名のスパンは $M=2.69$, $SD=0.70$, High WM・実験群6名のスパンは $M=2.00$, $SD=0.55$ であった。また、Low WM・統制群25名のスパンは $M=1.00$, $SD=0.00$, Low WMスパン・実験群5名のスパンは $M=1.00$, $SD=0.00$ であった。

この群分けの妥当性を確認するために、リスニングスパンテスト得点に関して、WM容量の大小（上位群，下位群）×教示（統制群，実験群）の 2×2 の分散分析を若齢者と高齢者ごとに行った。若齢者に関しては、WM容量の主効果（ $F(1, 94) = 220.27$, $MSE = 0.18$, $p < .05$, *partial eta squared* = 0.701）だけで有意差が認められた。教示の主効果（ $F(1, 94) = 0.78$, $MSE = 0.18$, *n.s.*, *partial eta squared* = 0.008）および両者の交互作用（ $F(1, 94) = 1.12$, $MSE = 0.18$, *n.s.*, *partial eta squared* = 0.012）はいずれも有意ではなかった。高齢者に関しても、同様の分散分析の結果、WM容量の主効果（ $F(1, 40) = 149.08$, $MSE = 0.12$, $p < .05$, *partial eta squared* = 0.788）だけで有意差が認められた。教示の主効果（ $F(1, 40) = 5.89$, $MSE = 0.12$, *n.s.*, *partial eta squared* = 0.128）および両者の交互作用（ $F(1, 40) = 7.16$, $MSE = 0.12$, *n.s.*, *partial eta squared* = 0.152）はいずれも有意ではなかった。したがって、高齢者の群の人数に極端な偏りがあるとはいえ、群分けは妥当であると判断した。

8. 日本語版TCAQに関しても、中央値分割によって上位群、下位群に分け、統制群、実験群別に調べた。若齢者の高TCAQ・統制群25名の得点は $M=72.92$, $SD=5.28$, 高TCAQ・実験群25名の得点は $M=72.68$, $SD=7.56$ であった。また、低TCAQ・統制群24名の得点は $M=56.54$, $SD=6.95$, 低TCAQ・実験群24名の得点は $M=54.11$, $SD=7.36$ であった。一方、高齢者の高TCAQ・統制群18名の得点は $M=80.83$, $SD=8.16$, 高TCAQ・実験群6名の得点は $M=72.33$, $SD=6.41$ であった。また、低TCAQ・統制群15名の得点は $M=59.67$, $SD=7.81$, 低TCAQ・実験群5名の得点は $M=57.40$, $SD=5.50$ であった。

ワーキングメモリ容量の分析と同様、これらについて、得点（上位群、下位群）×教示（統制群、実験群）の 2×2 の分散分析を行ったが、いずれの主効果も交互作用も有意ではなかった。すなわち、若齢者に関しては、TCAQ得点の主効果（ $F(1, 94) = 178.67$, $MSE = 46.74$, $p < .05$, $partial\ eta\ squared = 0.655$ ）だけで有意差が認められた。教示の主効果（ $F(1, 94) = 2.74$, $MSE = 46.74$, $n.s.$, $partial\ eta\ squared = 0.028$ ）および両者の交互作用（ $F(1, 94) = 2.28$, $MSE = 46.74$, $n.s.$, $partial\ eta\ squared = 0.024$ ）はいずれも有意ではなかった。高齢者についても、同様の分散分析の結果、TCAQ得点の主効果（ $F(1, 40) = 72.56$, $MSE = 57.81$, $p < .05$, $partial\ eta\ squared = 0.645$ ）に加え、教示の主効果（ $F(1, 40) = 4.59$, $MSE = 57.81$, $p < .05$, $partial\ eta\ squared = 0.103$ ）で有意差が認められた。両者の交互作用（ $F(1, 40) = 1.37$, $MSE = 57.81$, $n.s.$, $partial\ eta\ squared = 0.033$ ）は有意ではなかった。若齢者、高齢者それぞれの4つの群のTCAQと d' 値の相関係数を求めたところ、いずれの群でも有意ではなかったため、TCAQについては、これ以上の分析は行わない。

9. High WM・統制群22名のスパンは $M=3.64$, $SD=0.23$, 高WM・実験群27名のスパンは $M=3.72$, $SD=0.25$ であった。Low WM・統制群42名のスパンは $M=2.14$, $SD=0.48$, 低WM・実験群46名のスパンは $M=2.02$, $SD=0.60$ であった。このスパン得点に関して、実験2と同様、WM容量

の大小（上位群，下位群）×教示（統制群，実験群）の 2×2 の分散分析を行ったところ，WM容量の主効果だけが有意であった（ $F(1, 133) = 379.57, MSE = 0.21, p < .05, partial\ eta\ squared = 0.741$ ），教示の主効果 $F(1, 133) = 0.36, MSE = 0.21, n.s., partial\ eta\ squared = 0.003$ および両者の交互作用 $F(1, 133) = 1.57, MSE = 0.21, n.s., partial\ eta\ squared = 0.012$ はいずれも有意ではなかった。したがって，実験2の群分けは妥当であると判断した。

10. 日本語版TCAQに関しても，実験2と同様，中央値分割によって上位群，下位群に分け，統制群，実験群別に調べた。高TCAQ・統制群25名の得点は $M=73.74, SD=8.25$ ，高TCAQ・実験群58名の得点は $M=72.36, SD=7.03$ であった。また，低TCAQ・統制群51名の得点は $M=51.63, SD=7.47$ ，低TCAQ・実験群56名の得点は $M=52.80, SD=7.52$ であった。実験2と同様，これらについて，得点（上位群，下位群）×教示（統制群，実験群）の 2×2 の分散分析を行った結果，TCAQ得点の主効果（ $F(1, 215) = 412.23, MSE = 57.34, p < .05, partial\ eta\ squared = 0.657$ ）だけで有意差が認められた。教示の主効果（ $F(1, 215) = 0.016, MSE = 57.34, n.s., partial\ eta\ squared = 0.0001$ ）および両者の交互作用（ $F(1, 215) = 1.55, MSE = 57.34, n.s., partial\ eta\ squared = 0.007$ ）はいずれも有意ではなかった。実験2と同様，4つの群のTCAQと d' 値，メンタル・ローテーション・テストの正答数の相関係数を求めたところ，低TCAQ・実験群のメンタル・ローテーション・テストの正答数とTCAQで $r = -0.249$ （ $t = -1.71, df = 46, SE = 0.15, p < .10$ ），高TCAQ・統制群の d' 値とTCAQで $r = 0.661$ （ $t = 3.94, df = 22, SE = 0.22, p < .05$ ）であった以外は，いずれの群でも有意ではなかったため，TCAQについては，これ以上の分析は行わない。

11. すでに，梅本（1987）は，その著書『認知とパフォーマンス』において，記憶パフォーマンスも含めたステレオタイプ脅威と類似の現象に関して，実証的知見に基づく論考を加えている。