

The Einstellung effect in anagram problem solving: evidence from eye movements.

Jessica J. Ellis & Eyal M. Reingold, *Frontiers in psychology*, 2014,(5),679,1-7

Introduction

- 刺激に対する親密性や獲得済みの先行知識が問題解決のパフォーマンスを損なうという概念は、様々な関連する用語によって言及されてきた。
 - cf) 機能的固着 (functional fixedness), ネガティブ転移 (negative transefer), 心的構え (set effect), Einstellung 効果 (Einstellung effect)
- 機能的固着とはオブジェクトの習慣的な使用法がほかの用途での使用を妨げるような場合をいう。
 - 有名な Duncker(1945)の”ロウソク台課題”では、マッチを入れる容器として提示されたことで、それ以外の使用法(ロウソク台にする)を妨げるものである。
- ネガティブ転移は以前に獲得した刺激-反応間の連合が、新しい刺激と反応の獲得及び維持を阻害するものである(e.g., Schultz, 1960; Sweller, 1980; Chrysikou and Weisberg, 2005; Landrum, 2005; Osman, 2008).
- 問題解決セットまたはメンタルセットは、事前に行われた課題によるネガティブな効果のことで、それはよく知っているが最適ではない解法を引き出し、他の解決策が検討されないようになる効果。
 - Luchins(1942)の一連の水がめに関する実験によって明らかにされた Einstellung effect はメンタルセットのネガティブな効果の優れた実例の1つである。
 - このパラダイムでは、似たような解法を持つ複数の問題にさらされることによって、問題解決において常習的なアプローチをとるように誘導され、その後、習慣的な解決策が適切ではないような状況にさらされると、多くの参加者が問題が解決できないと主張する。
 - しかし、事前の問題に取り組んでいない参加者は、すぐに解法を見つけることが出来る。
 - つまり、解決者が経験する困難感(次善の経験の悪影響)を表している。
- エキスパート研究の知見を考えたときに、過去の経験や刺激に対する慣れが問題解決のパフォーマンスを妨げるかもしれないという研究は、相反しているように思われる。
 - しかし、専門家も次善の経験や刺激によるネガティブな効果とは無縁ではないことは、医療分野の研究、チェスを使用した研究で示されている。
 - 例えば Bilalić et al. (2008a)は、チェスの専門家の Einstellung effect を検討するために、眼球運動を用いて検討した。
 - この課題では、プレイヤーは最も短い手数でチェックメイトをすることを求められた。この盤面では2つの可能な解決策が存在した。良く知られている5手でチェックメイトをする解法(次善解)、あまり知られていない3手(最善解)でチェックメイトを行うことが出来る解法であった。
 - 次善解を発見した後も参加者は、最適な解を探していると報告した。しかし、眼球運動は次善解のある領域に最善解のある領域よりも多くとどまっていた。
- このように Einstellung effect 効果では、記憶内のスキーマの活性によってそれ自体に関する注意が促され、他の情報に対する注意がそらされることが示された。

- 本研究では、刺激への習熟が問題解決に及ぼす悪影響をさらに調べるために、Ellis et al.(2011)によって導入されたアナグラム課題を実行しながら、参加者の眼球運動を監視した。
 - アナグラム課題はほとんどの大人が持つ言語という専門知識(言葉の親密性)を利用して、Einstellung effect の効果を検証する機会を与える。
 - この課題が他の Einstellung effect の課題より優れている点は、この課題は一度の実験で様々な構えの効果のバリエーションを試せることにある。
- アナグラム課題は洞察の研究や問題解決におけるメンタルセットのネガティブな効果の影響を示すために長く使用されてきた。
 - 目的語が単語(HEART)の時に、元の単語が非単語(THREA)よりも単語(EARTH)の時のほうが時間がかかり、解決率も下がることが示されている。
- 本研究では、アナグラム解決課題を使用して眼球運動を観察した。
 - 図 1 に示すように本研究におけるアナグラム課題では、中央に位置する 3 つの文字列と、周囲に独立して表示される 3 つの文字の合計 6 つの文字で構成されていた。
 - 中心の 3 つの文字列は単語条件と、非単語条件
 - 参加者は 5 文字の単語を作成するように求められ、6 つの文字のうちの 1 つはディストラクターで、完成された文字に使う必要はないことが教示された。
- 似た課題を使用した Ellis et al(2011)では、トライアルの開始時点ではディストラクターと他の単語の間の区別はされず、同様に探索の対象になっていたことが示されている。
 - そして思考の終わりに向かって、ディストラクターの注視時間は、正解の単語に比べて減少し、この変化は解答の数秒前に起こった。
 - この変化、参加者が回答の単語に対する主観的な洞察経験を報告をしたかに関わらず起こった。
 - この結果は、洞察の主観的報告と知識の蓄積の間の乖離を示すものである。
- 本研究では、Ellis らの結果の再現と、中心の単語がよく知ったものであるときよりも、知らないものであるほうが、解決時間が短く、回答率も高くなることを検討する。
- そして最後に、これらの親密性によって視線運動がどのように変化するかについて検討する。

Materials and Methods

participants 60 人の学生で \$10 のクレジットを支払った。すべての参加者が流暢な英語話者であった。

Apparatus

- SE Research EyeLink 1000 を使用して参加者の眼球運動を記録した。
- サンプリングレイトは 1000Hz、刺激は 75Hz のリフレッシュレート及び、1024×768 のスクリーン解像度を持つ 19 インチのモニター上に提示された。
- 参加者は画面から 60cm のところに座り、頭の動きを最小限にするためにあご台を使用した。
- フィグゼーション後注視位置の誤差は 0.5° 以下であった。

Materials

アナグラム課題の設定

- アナグラム課題は6つの大文字で構成されており、中央に位置する3文字の文字列と、中央の文字列の上と左右に表示された3つの個別の文字からなっていた。
- 課題はこの6つの文字を使用して5文字の文字列を作成することで
- 中央の文字列はすべて解凍語に属していた。
- 3つの単独提示された文字の中で2文字が解答語になるようになっていた。
- それぞれのアナグラム課題において解答語は1つだけでした。
- このことはディストラクターをしようして他の単語が作成できないことを保証する。

単語の親密性の操作

- 各アナグラム課題において、中心の3文字が単語か非単語かのどちらかで提示がされました。
- 単語条件では中央の3つの文字は意味を持つ単語になっており、非単語条件では中心の単語はシャッフルされ単語にはならないようになっていた。
- それぞれの条件で使用された文字は同じになるようになっていた。つまり、単語非単語条件の間の差異は、中央の文字列の並び順だけであった。
- 中心の単語の使用頻度は Brysbaert and New (2009 より、約 435/100 万(SD=1243)程度になるように統制されていた。

解答語

- 解答語は5つの文字で構成され、常に子音で始まり問題のうち33%が1つの母音を含むもので、66%が2つの母音を含んでいた。
- 解決語は Brysbaert and New (2009)より、使用頻度の平均は約 175/100 万(SD = 396)
- 中心の3つの単語は2つの子音と1つの母音から構成されていた。
- 参加者には中心の3つの語は常に含む必要があることを思い出させるために、色は緑色で表示され、黒で書かれたほかの3つの独立した文字よりも若干小さく異なるフォントで書かれていた。

Procedure

- 参加者は6回の練習試行と、72回の本試行を行った。
 - アナグラム課題の半分が「単語条件」、半分が「非単語条件」であった。
 - すべての課題は、画面中央に注視点が表示される画面から始まりました。
 - 1000ms 後、応答が行われるか45秒の制限時間が過ぎるまで課題が表示され、画面上に残りました。
 - 参加者には、反応の速さ最も重要であり、時々間違った答えをするかもしれないとしても、反応の前に解法を検証することをしないように勧めた。
 - 参加者は応答のためにボタンを押した。その後刺激が消え、参加者は実験者に回答を報告した。その後参加者は自分の回答が正しいかどうかのフィードバックを受けた。
 - 各施行の後、参加者はアナグラム課題を解決した際の主観的経験の報告を求められた。
- 1「解決語はどこから突然出てきた。私は答えを得るために何かをしたことに対して気づいていない」
 - 2「アナグラムを解決するために様々な文字を組み合わせましたが、どれもうまくいかなかった。そんなとき解決策が突然心に浮かんだ」

3「アナグラムを解決するために様々な文字を組み合わせました。私は段階的に解決策を立てることができました」

4「私はアナグラムを解決しなかった」

1, 2 …ポップアウト, 3 …非ポップアウト

Results

- この実験の目的は、中央の文字列のタイプが平均タスクパフォーマンスと眼球運動に与える影響を検討することだった。
- まず、中央の文字列に対して操作したことで Ellis(2011)の結果との差異がないかを検討する。
- そのために、まず単語-非単語間の差異を検討する、そのご中心の単語の親密性と参加者の主観報告の関係についても検討する。

Table 1 | Mean problem solving and eye movement measures shown overall for correct trials, and separately for word and nonword trials.

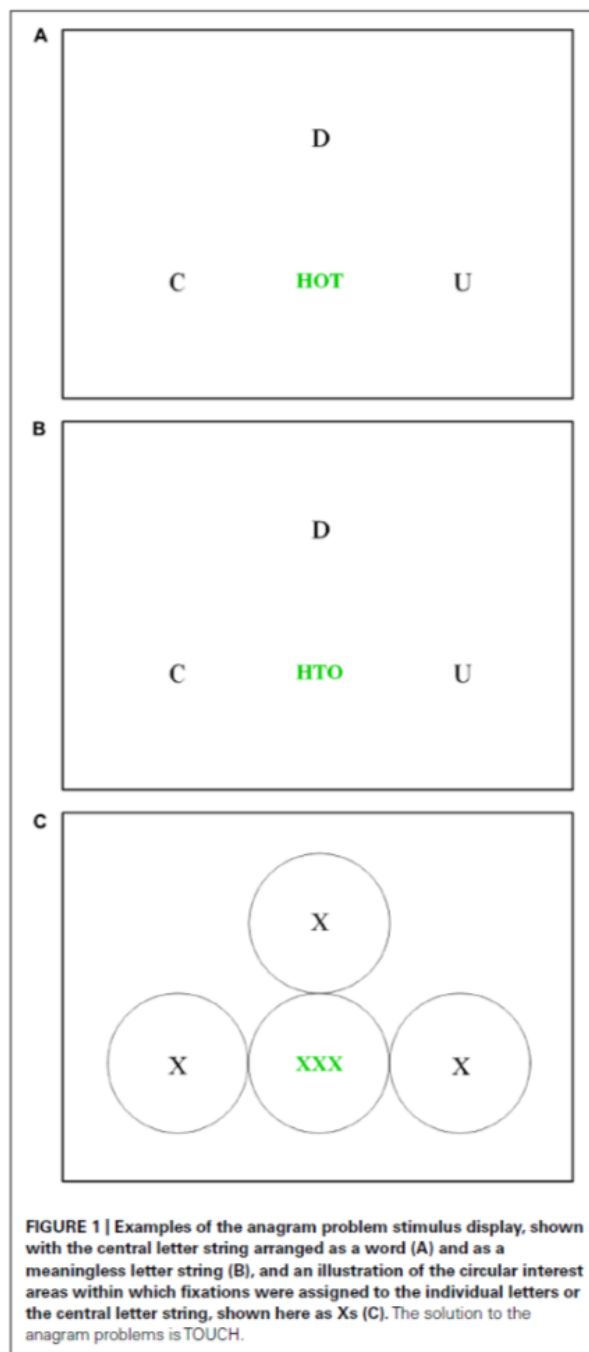
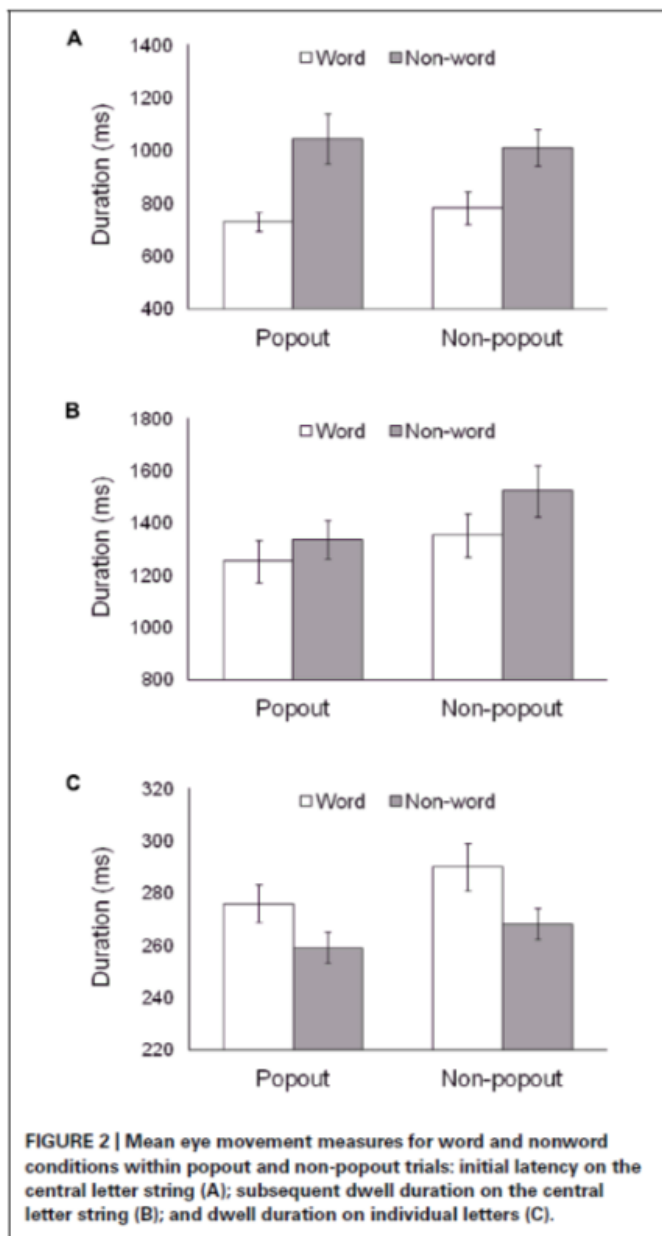
Variable	Overall	Word	Nonword	Significance
Accuracy (% correct)	52.7 (2.4)	51.7 (2.4)	53.8 (2.6)	$t(59) = 1.52, n.s.$
Response time (s)	14.3 (0.4)	15.3 (0.6)	13.4 (0.4)	$t(59) = 3.25, p < 0.01$
Number of dwells, central	7.39 (0.37)	8.11 (0.41)	6.76 (0.37)	$t(59) = 4.69, p < 0.001$
Number of dwells, letters	9.40 (0.55)	10.45 (0.63)	8.53 (0.53)	$t(59) = 4.80, p < 0.001$
Percentage of trials classified as popout	69.0 (3.4)	68.9 (3.5)	69.2 (3.6)	$t < 1, n.s.$

SEs are shown in parentheses.

- 52.7%が正答、試行の 4.2%が誤答で、43.1%は解決されずに時間切れとなりました。
- 正答試行の平均反応時間は 14.3 秒 (SD = 3.3 秒) であり、69.0%の試行は参加者によってポップアウトとして分類され、31.0%は非ポップアウトとして分類された。
- 単語の試行では非単語の試行に比べて反応時間はかなり遅かった
- 単語の試行における正答率は非単語の場合よりも低くなっているが、ここには有意な差が見られなかった。
- 最後に単語の親密性と、洞察経験の主観的報告との間に有意な差異は見られませんでした。

視線の分析(t 検定)

- 視線の分析は正しい試行のみでおこなわれ、特定の刺激に対する Fixation のみを対象とした。
- 具体的には個別の文字または中央の文字列から 192 ピクセルの円内にある場合を対象とした。
- 正答した参加者の中の 69.9%の Fixation は中心の単語に、28.3%は 3つの文字のうちのどれか1つに、残りの 1.8%は文字の周りの領域の外にあった。
- 割り当てられた Fixation は停留(dwelling)に変換された。停留は同じ領域内の1つの連続的な Fixation として定義された。
- Table1 に示すように、単語試行の応答時間が、非単語試行のより遅いのに対応して、非単語試行と比較して単語試行において、中央文字列と個々の文字に対する停留割合は多くなっていた。
- この結果は親密性が問題解決を妨げるという結果を再現した。



- しかし、図 2 に示すように、いくつかの眼球運動は中央の文字列のタイプによって影響を受けた。具体的には、

- (a) 中心の文字列に対する最初の滞在時間
- (b) 再び中心の文字列の領域に滞在した際の停留時間
- (c) 他の文字列における停留時間

それぞれの眼球運動の測度に対して、2(ポップアウト, 非ポップアウト)×2(単語, 非単語)の ANOVA を行った(Figure.2).

- その結果、ポップアウトの効果、交互作用は見られなかった。このことは主観報告の違いが、他の眼球運動の指標に対して影響を与えないことを示している。
- (a)中心の文字列に対する最初の滞在時間は、単語の試行では非単語試行よりもはるかに短かった (Figure 2A; $F(1,59) = 29.78, p < 0.001$)。
- (b)その後の中央文字列の停留においてもこの傾向は維持された (Figure 2B; $F(1,59) = 7.78, p < 0.01$)
 - この単語試行のアドバンテージはワーキングメモリに保持する時間のアドバンテージだと考えられる。
- (c)逆に個々の単語の停留時間は単語試行に比べて非単語試行においてより長い停留時間がかかっていた (Figure 2C; $F(1,59) = 25.31, p < 0.001$)。
- この停留時間の差異は単語試行において中央文字列のもつ単一のゲシュタルトに対して、個々の文字を中央の文字列に統合するのが困難になるためだと考えられる。

課題の前半後半の眼球運動の差異

- 加えて、ディストラクターと解答語を見ている時間を、課題の前半後半で分けて、測定した。そして、注視の単語のタイプと主観報告タイプの2つの要因から検討した。
- Ellis ら(2011)に基づいて考えると、ディストラクター文字と解答文字の時間は試行の開始時には同じであると考えられるが、後半ではディストラクターへの注視時間を減少させると予測される。
- 本研究ではそれぞれの試行に対して、ディストラクター文字に費やされた時間の割合と2つの解答文字の平均を別々に比較した。
- 独立変数として、単語タイプ(ディストラクターvs 解答語)、主観報告(ポップアウト vs 非ポップアウト)、中央文字列タイプ(単語 vs 非単語)を従属変数としてそれぞれの領域の停留割合とし、 $2 \times 2 \times 2$ のANOVAを行った

前半の分析

- Figure.3のA~Dに見られるように、すべての条件において前半試行ではディストラクターと解答語の間に有意な差異は見られませんでした。
- またANOVAにおいても主効果も交互作用は見られなかった。

後半の分析

- 対して、後半のすべての条件において、ディストラクターと比較して解答語のほうが停留期間が長かった ($all ts > 3.83, all ps < 0.001$)。
- ANOVAは文字タイプの効果見られたが ($F(1,59) = 75.65, p < 0.001$)他のどの効果も見られなかった ($all Fs < 0.23, n.s.$)。
- 中心語による効果が出なかったことから、洞察に先立つ知識蓄積の効果が、親密度の効果と独立して起こっていることを示している。

解決できなかった参加者の分析

- 最後に解決できなかった場合についても検討した (Figure.3, EF)。
- アナグラムの失敗に関しては、単語タイプの違い(ディストラクターと解答語)の効果のみ見られた ($F(1,59) = 5.51, p < 0.05$)、他の主効果やインタラクションは見られなかった ($all Fs < 2.48, n.s.$)

Discussion

- 本研究の主な目的は、簡略化されたアナグラム課題で親密度がパフォーマンスに及ぼすネガティブな効果を検証することであった。
- アナグラム課題においてスクランブル文字として提示するときよりも単語形式で提示した時のほうがパフォーマンスが悪くなるという先行研究の結果を追試した。
 - この効果は、文字の順序を並び替えて形成する際に、完成されたゲシュタルトを壊すことが難しいためだと考えられている。
- しかし、本研究の参加者の眼球運動は、刺激習熟がアナグラムの問題解決に及ぼす影響のより複雑なパターンを示した。
 - 具体的には、本研究では、中央の文字列が非単語の時よりも単語形式で提示されたときの停留時間が短くなることを示した。
 - このことは、中央の文字列が意味のある単語の時に意味のない単語の時に比べて、よりワーキングメモリにエンコード・維持しやすかったためだと考えられる。
- さらに興味深いのは、眼球運動から親密性がアナグラム課題に影響を与えるかを示した点
 - 中央の文字列が単語として提示されたときに、刺激に対する親密性による干渉が個別の文字に対する長い注視時間に起因することを示した。
 - ゲシュタルト心理学によって説明されるように、個々の文字に対するながい注視時間は、完成された単語に個々の文字を組み込むことが困難だったことを示している。
- *Einstellung effect* に関する以前の眼球運動を用いた研究では、特定のスキーマをワーキングメモリに活性化させると、そのスキーマに対する知覚的偏りが生じ、珍しい最善解を見つけるために必要な情報から離れてしまうことが示された。
 - しかし本研究は、中心の単語が意味を持つものであるときに、その単語に対する知覚的なバイアス(視線の集中)を示さなかった。
 - この結果は、中心の単語が親密性の高い者であることがアナグラム課題においてアドバンテージになっていたためであると考えられる。
 - このアドバンテージによって、参加者は比較的中心を見ることなく情報を使用してアナグラム課題を解くことが出来たのだろう
- 本研究の結果やこれまでの研究は、領域知識や刺激に対する親密性はエキスパートのパフォーマンスの基礎となるものであるが、柔軟性を要求されるような問題解決状況においては解決の妨げになることを示している。

