

Hedne, M. R., Norman, E., & Metcalfe, J. (2016). Intuitive feelings of warmth and confidence in insight and noninsight problem solving of magic tricks. *Frontiers in psychology*, 7, 1314.

洞察問題解決における自動/無意識と制御/意識的プロセスの関与

- 進捗モニタリング理論/満足進捗理論(progress monitoring theory/satisfaction progress theory)
 - ✓ 問題解決は、問題解決行動の意識的な、ステップバイステップのモニタリングを含む
 - ✓ このモニタリングは問題解決行動の心的シミュレーションと、その成功失敗の評価によって行われる。
- ⇒問題解決は意識的であること、意図的な計画、監視および評価を伴う
- 表象変化理論(Ohlsson, 1992; Knöblich et al., 1999)
 - ✓ 洞察の問題解決は最初に誤った問題空間の構築を含む
 - ✓ 表象の変化は、制約の緩和(不必要な仮定の消去, チャンク分解)を通して行われる
 - ✓ この理論では誤った問題表象やそれを解決するメカニズムも意識的な精査をする必要はない
- ⇒自動化された無意識のプロセスによって特徴付けられる

洞察のプロセスについての議論

- 突然/不連続的なプロセスとしての洞察
 - ✓ 洞察は意識的な分析的思考とは独立した非意図的な無意識の処理の産物であることを暗示(Smith and Kounios, 1996)
 - 徐々に/連続的なプロセスとしての洞察
 - ✓ 洞察は緩やかでより連続的なプロセスの結果である
 - ✓ 問題解決の試みの過程で、問題表現が無意識的/曖昧なものから意識的/言語化可能なものへと変化していく
- ⇒いずれの見解も、洞察の主観的な経験は、関連する無意識/潜在知の活性化を伴う
- ⇒洞察には潜在的・無意識的な知識が必要であることが一般的に合意されている

直観と洞察の関係

主観的感情と無意識の知識との関係は、潜在学習を含む他の形態の潜在認知に関連して広く研究されてきた (例: Dienes and Scott, 2005; Norman and Price, 2015)

- Metcalfe and Wiebe (1987)の知見
 - ✓ 参加者に、問題に取り組んでいる間、一定の間隔で暖かさの評価(WR)を提供するように求めた
 - ✓ 多段階問題/パズルと洞察課題の間の暖かさの評価の差異を検討
 - ✓ WR は多段階問題の正解を出す前には徐々に上昇していくが、洞察問題の正解を出す前にはあまり上昇しない
 - ⇒洞察と非洞察の間の質的な差異
 - ⇒主観的な感情(WR)が問題解決中の知識の利用と関与
- 問題点:そもそも使用している課題が異なるため、行動や主観的報告の違いは課題の構造に依存している可能性がある

D2 二宮

➤ Danek et al. (2013, 2014a,b) の知見

- ✓ これらの制限に対抗するために新しい実験パラダイムを開発
 - ✓ 手品の動画を見せ、そのトリックを考えるという課題
 - ✓ 手品のトリックは、時として制約緩和の結果として洞察的に解決することもあり、また、体系的に異なる可能性を検討することを伴うようなステップバイステップの方法で解決されることもある (Danek et al., 2014a)
 - ✓ 実験では解決策が洞察の経験と関連しているかどうかについて報告させた
 - ✓ 洞察的な解決は、非洞察な解決よりも正確性の高さ、確信度の高さ、動画提示の回数と関連
- ⇒ 洞察問題解決は、非洞察問題解決とは質的に異なるという考えを裏付ける
- ⇒ この方法はタスクの違いによる交絡要因を排除して、2 種類のタイプの問題解決における直観的な感情 (WR や確信度) と客観的な指標 (正答率など) との間の関係を探ることが可能になる

*以降では問題解決の段階的なステップに関する仮定ではなく、Aha!のような主観的な経験によって特徴づけられる問題解決を「洞察問題解決」と呼ぶ

➤ Danek らの研究の制限

1. 洞察の定義に、確信度に関する基準が含まれていたこと

- ✓ 参加者は、「Aha!」の経験は、「自分の解答が正しいと比較的確信している」と感じることで特徴づけられると説明される (p.662)
 - ✓ 本研究の関心は、確信度のような直感的な感情と、実際のパフォーマンスとの関連であり、このような説明は評価に影響を与える可能性がある
- ⇒ 何が洞察解を構成するかについて参加者に与えた定義に確信度に関する情報が含まれないようにした。

2. 解答時間の指標の正答率が低さ

- ✓ 絶対的な解答時間を秒単位ではなく、動画提示の回数 (1~3 回) で検討したため、正確ではない (しかも動画ごとの時間が異なる)
- ⇒ 各動画の時間を標準化し、解答時間の測定値としてミリ秒を用いた 1。

3. 確信度と正答率の関係を体系的に評価していない

⇒ 洞察と非洞察のそれぞれについて、確信度と正答率の関係を検討

4. 直感的感情の尺度は回顧的確信度であり、解答試行中の直感的感情の測定は含まれていなかった

⇒ 解答が与えられる前の解答への近さの直感的感情 (暖かさ) を評価

5. 解答が与えられた後の安定性の尺度を持っていなかった。

- ✓ 洞察のほうがより結晶化された構造として思いつく
- ✓ 洞察のある解決策が非洞察のある解決策よりも安定していたとしたら、人々はそれを変えることはほとんどないだろうと予想される

⇒ 提案された解決策に固執する傾向を、可能な解決策の選択肢をいくつか与える多肢選択課題を行うことで測定

本研究の目的

目的 解決すべき問題が手品のようなものである場合に、洞察的解法と非洞察的解法の間で、直感的な感情(暖かさの評価, 確信度)と行動指標(正答率, 反応時間, 提案した解法の変更)の関係を検討
⇒問題解決の2つの形態における知識の意識的な利用可能性についての理解に貢献する

【二宮補足

この研究目的は、大きく以下の3点です。

①洞察と非洞察の間の質的な差異の存在を検討すること

*この観点についてはすべての指標で洞察と非洞察の間の差異があるかという点を検討します

②洞察における直感的な感情の1つである確信度と正答率の間の関係について議論することで、洞察における無意識的知識の利用についての議論をすること

*確信度と正答率が対応していない場合、問題解決についての知識にアクセスできていなかったということになり、無意識的性質を強調する。

*WR が回答直前でも上昇しないことは、洞察の突然性や問題解決の知識に意識的にアクセスできないことを示す

③手品を使用した方法論のアップデート

】

方法

参加者 ベルゲン大学（人文科学部，法学部，数学・自然科学部，医歯学部，心理学部）から19～31歳（ $M = 21.81$, $SD = 2.55$ ）の学生51名（男性14名，女性37名）を募集

手品刺激 手品・トリックは、最初の状況，手品・モーメント，啓示（de Ascanio, 1964/2005）から構成されており，15秒という時間枠の中でこれら3つのフェーズ含んでいた。

- ✓ すべてのトリックが，固有の方法で行われる
- ✓ 使用されたすべての手順は，比較的少ない言葉で簡単に説明できるもの
- ✓ それぞれの手品は，ミスディレクションや余計なジェスチャーをほとんど使わずに，問題解決の課題として提示された
- ✓ 選ばれた32の手品のうち，20の手品はDanekら(2014b)の研究と同様
- ✓ 32試行のそれぞれで，プロのマジシャンが手品を披露する動画が提示
- ✓ 各動画クリップは15秒の長さ
- ✓ 手品のトリックには観客からの注意を集めることが必要なため，13本はマジシャンが直立している様子を撮影したもの，6本はテーブルの後ろに立っている様子を撮影したもの，13本はマジシャンの手とテーブルの上の様子を撮影したものであった。

例1 動画 https://www.youtube.com/watch?v=_jE25LbLaoQ

例2 動画 <https://www.youtube.com/watch?v=YTvTFNnwDEg>

例3 動画 <https://www.youtube.com/watch?v=VqNYrADykUk>

手続き

教示

- ✓ 実験の開始時に、参加者には、全体的な手順に関連した口頭での指示に加え、「Aha!体験」の定義も与えられた。

Aha!体験の定義

- ◇ 「どこからともなく」現れた解決策であり、他の解決策や以前に提案された解決策とは異なるものである
- ◇ その解決策に至るまでの推論プロセス全体を説明することができれば、これは「Aha!体験」とはみなされない

*ここで重要なのは Danek et al(2013, 2014a,b)とは異なり、確信度に関する言及を含まないこと

問題解決タスク

32の手品の動画を見て、そのトリックを答える課題

➤ 各トライアルの構成 (figure 4)

- ✓ WR は、青 ("冷たい") から赤 ("暖かい") のグラデーションで着色されたバーからなる視覚的アナログスケール(VAS)上をマウスでクリックして報告. 応答が与えられなかった場合、WR スケールは 4000ms 後に消えた
- ✓ 参加者は、表示されている手品の解答がわかったら、スペースを押すように指示された
- ✓ スペースを押すと、進行中の試行が中断され、最初の動画提示以降は、どの時点でも中断可能
- ✓ 参加者がスペースを押さなかった場合、シーケンスは前述の最大持続時間まで終了した
- ✓ 参加者は、スペースを押し中断した場合も、タイムアウトで終了した場合も、"Aha!体験しましたか?" という質問が提示され、参加者は「はい」または「いいえ」を示すことでこれに答えた
- ✓ その後、画面上にテキストボックスが現れ、そこに手品の解答を入力するか、トリックがどのように行われたかの仮説がない場合は「わからない」と記入した

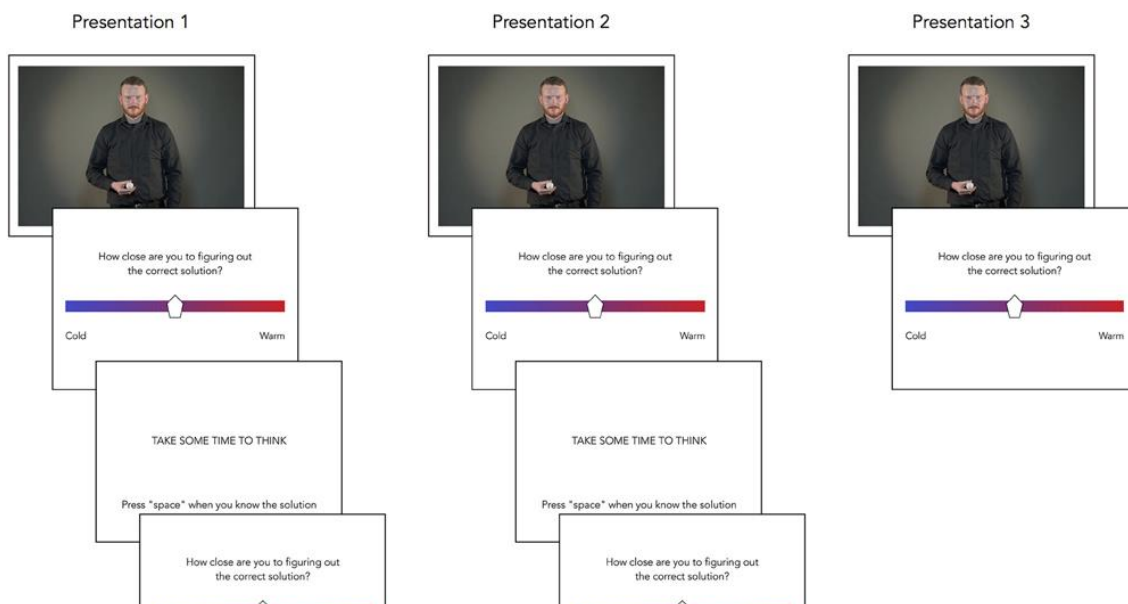


Figure 4 各試行のながれ

D2 二宮

➤ 確信度の測定

- ✓ 回答の入力後、確信度を視覚的スケール(薄い灰色(「全く確信度がない」)から濃い灰色(「完全に確信度がある」)までのグラデーション)で回答

➤ 多肢選択課題

- ✓ 確信度を報告した後、参加者は4つ手品のトリックに関する多肢選択課題を与えられた。
- ✓ そのうち1つが正しいものであった。
- ✓ 選択を行った後、その選択に対する確信度を評価
- ✓ その後、多肢選択を回答する際の戦略を、「4つの選択肢をすべて見て比較した後、最も可能性が高いと思ったものを選んだ」、「選択肢の1つを見た瞬間、それが正しいものでなければならないと思った」、「私が選んだ選択肢は、私の書いた解答に最も近いものであった」、「すべての選択肢の中で同じように不確実性を感じ、無作為に1つを選んだ」
- ✓ 『各試行(Tig 4) → 確信度評価 → 多肢選択課題』の流れを32回繰り返した。

事後アンケート

32回の試行後、趣味で手品をやっていた人を知っているか、あるいは自分自身が手品をやっていたことがあるか、という質問と、一般的な手品の知識を超えた手品の知識を持っているかどうかを聞いた。

結果

測定項目

- ✓ 始めに、手品のトリックがわからないと回答したり、反応のない施行を除外した。
- ✓ 回答の正確性は2人の評価者(どちらもプロのマジシャン)が、4段階(完全に不正解-1, 大部分が不正解-2, 大部分が正解-3, 完全に正解-4)の尺度で採点した
 - ◇ Cronbachの α 係数を用いて測定した評価者間信頼性は0.911であった
 - ◇ 回答の正答率の値として、上記の評価を二分したものを使用する。
1または2=不正解, 3または4=正解
- ✓ 反応時間はミリ秒単位で測定
- ✓ 暖かさと確信度は1から100までの整数値で測定した。

データのフィルタリング

- ✓ 複数の解答が与えられた場合、参加者が手順を中断しなかった場合を除外
- ✓ 参加者のうち8人のデータは、Aha!体験を一度も報告しなかったため、完全に除外した
- ✓ 手品の1つを含む試行は、正解を報告した人がいなかったため、全参加者から除外
- ✓ 参加者が回答や実験後のアンケートで、手品がどのように行われるかについて事前に知っていたと報告していた場合
⇒上記の基準を満たさない試行(661件)を除外、残り971件の試行を解析した

洞察 vs 非洞察課題

➤ 記述統計

- ✓ 対象となった試行(N=971)のうち、29%は洞察を使って解いたと報告した。

D2 二宮

- ✓ 洞察の有無に関わらず、トリックが解かれる頻度にはかなりのばらつきがあった。

例：例1 動画は、洞察を使って最も頻繁に解かれた問題

例3 動画は、洞察を使って解く頻度が最も低い問題

➤ 統計解析

- ✓ 自己申告した解答タイプ（洞察と非洞察）を独立変数として、正答率（正解/不正解）、回答時間を従属変数として、t検定をそれぞれ実施
 - ◇ 洞察試行（ $n = 281$, $M = 0.57$, $SD = 0.50$ ）と非洞察試行（ $n = 690$, $M = 0.37$, $SD = 0.48$ ）では、各タスクの正答率に有意な差がありました； $t(506.9) = 5.78$, $p < 0.001$, $d = 0.51$.
 - ◇ 洞察試行（ $M = 38.23$, $SD = 18.69$ ）と非洞察試行（ $M = 40.56$, $SD = 19.53$ ）の間に回答時間に、有意な差異は見られなかった； $t(969) = 1.705$, $p = 0.089$, $d = 0.10$
 - *ただし、片側t検定で有意になる（ $p < 0.05$ ）

暖かさの評価

➤ 評価値について

- ✓ 洞察試行と非洞察試行の暖かさ評価の時間的な変化を比較した解析では、3または4回のWR測定を行った試行のみに対して行った。

【補足：そうしないと後述する指標が計算できないから。これにより70%ほどの試行を除外】

- ✓ 暖かさの評価には2つの指標を用いた。

1. 差分WR 最初のWRから、回答を中断する前の最後のWRの差分を求めた
2. 角度(angular)WR 差分WRを秒(おそらく、差分時間)で割る

⇒これらの値が大きいことは、徐々に解に近づいた感覚があったということである。

⇒洞察に関連していない試行では、微分暖かさと角度暖かさの評価に高い値が見られると予想【補足：洞察は解けるまでWRは上がらないのに対し、非洞察は徐々に上がっていきと考えられるから】

➤ 統計的検定

- ✓ 洞察試行（ $n = 50$, $M = 2.92$, $SD = 23.26$ ）と非洞察試行（ $n = 162$, $M = 0.30$, $SD = 19.52$ ）の間で、差分WRに有意差なし； $t(210) = 0.79$, $p = 0.429$, $d = 0.12$.
- ✓ 洞察試行（ $M = .41$, $SD = 2.60$ ）と非洞察試行（ $M = -0.004$, $SD = 0.29$ ）試行の間には、角度WRの評価に有意差なし； $t(49.36) = 1.12$, $p = 0.266$, $d = 0.22$

確信度

➤ 洞察と非洞察の間の差異

- ✓ 洞察試行（ $M = 78.32$, $SD = 20.35$ ）と非洞察試行（ $M = 68.95$, $SD = 23.96$ ）の間には、確信度に有意な差があった； $t(606.8) = 6.17$, $p < 0.001$, $d = 0.50$

⇒洞察の教示に回答への自信についての記述を入れずのこの差異が出たことが重要。

【補足：ここまでは洞察と非洞察の間差異を強調する結果】

➤ 確信度と回答の正確性の関係

- ✓ 2つの解法にける、確信度と正答率の関係を個別に比較するために、2つの分析を実施した

D2 二宮

1. 正答率を独立変数とし、それぞれの解のタイプについて t 検定を実施
 - ◇ 洞察解では、正解の確信度 ($n = 160, M = 81.34, SD = 17.16$) は、不正解の確信度 ($n = 121, M = 74.32, SD = 23.41$) よりも有意に高かった; $t(210.97) = 2.78, p < 0.01, d = 0.34$
 - ◇ 非暗示解では、正解の確信度 ($n = 254, M = 74.31, SD = 22.73$) は不正解の確信度 ($n = 436, M = 65.83, SD = 24.14$) より優位に高かった; $t(688) = 4.55, p < 0.001, d = 0.36$.
 2. 2 つの条件における確信度と正答率の関係を、信号検出理論 (SDT) 統計量 Az (Macmillan and Creelman, 2004; Norman and Price, 2015) を用いて比較
 - ◇ この値は、「あるレベルの確信度が与えられたときの回答が正しい確率」を表し、個人のメタ認知能力を示す指標 (Song et al., 2011, p. 1789)
 - ◇ Az スコアが 1 の場合、正解と不正解の区別が完全にできていることを示し、 Az スコアが 0.5 の場合はランダムな回答を示す。
 - ◇ Az スコアは個々の被験者ごと洞察試行、非洞察試行について計算された
 - ◇ 有効な Az スコアを持っていた 33 人の参加者の洞察試行 ($M = 0.56, SD = 0.28$) と非洞察試行 ($M = 0.63, SD = 0.18$) における Az スコアを比較した
 - 両群間に有意な差はなした $t(32) = 1.06, p = 0.297, d = 0.30$
 - 洞察試行の平均 Az スコアには、無作為回答 (0.5) との有意差はなかった ($M = 0.57, SD = 0.28$); $t(33) = 1.51, p = 0.142, d = 0.25$
 - 非洞察試行では、平均 Az スコアは無作為回答 (0.5) よりも有意に高かった ($M = 0.64, SD = 0.17$); $t(41) = 5.43, p < 0.001, d = 0.82$.
- ⇒ 洞察的に問題解決した場合には、自分が正しいか間違っているかを判断することができないが、非洞察的に問題解決をした人はこの区別が可能だった

多肢選択課題

- 統計的分析
- ✓ 回答を保持するかを検討するために、参加者が自分の解答に最も類似した選択肢を選んだと解釈した試行と、それ以外の戦略を用いて選択した試行に試行を分けた
 - ✓ これを二項変数として再コード化した (「私が選んだ選択肢は、私の書いた解答に最も似ているもの」 -1 ; 「その他の戦略」 -0)
 - ✓ そして、洞察と非洞察の間でこの変数を t 検定
 - ◇ 洞察試行 ($M = 0.72, SD = 0.45$) と非洞察試行 ($M = 0.61, SD = 0.49$) の間に有意な差; $t(969) = 3.37, p = 0.001, d = 0.23$
 - ◇ 回答が正しい試行のみで同様の分析をすると、洞察試行 ($M = 0.79, SD = 0.41$) と非洞察試行 ($M = 0.78, SD = 0.41$) の間に有意差なし; $t(412) = 0.25, p = 0.80, d = 0.02$
 - ◇ 回答が正しくない試行のみで同様の分析をすると、洞察試行 ($M = 0.63, SD = 0.49$) と非洞察試行 ($M = 0.51, SD = 0.50$) の間に有意な差があった; $t(196.58) = 2.41, p = 0.017, d = 0.24$
- ⇒ 参加者は非洞察よりも洞察において、自身の不正解の回答を保持する傾向が強いことが示された。

考察

目的 解決すべき問題が手品のようなものである場合に、問題解決のためのメタ認知的な「直感的な」感情と客観的な指標との関係が、洞察と非洞察では異なるかを検討することで、問題解決における意識的なプロセスと無意識的なプロセスに関する議論に貢献すること

【洞察と非洞察の間の差異についての議論】

➤ 正答率と解決時間

- ✓ 先行研究(Danek et al., 2014b)と同様に、洞察的解は非洞察よりも正答率が高かった
- ✓ 一方、回答時間に関しては、先行研究とは異なり、2つの解法間で差異はなかった
⇒先行研究の動画時間の長さの差異が大きかったことに起因

➤ 暖かさの評価

- ✓ 予測に反して、洞察のある解決策と非洞察のある解決策の暖かさの評価の変化に差はなかった
先行研究と異なる結果がみられた理由

- ◇ 2つのタイプの解決策には、同じ根本的な問題解決プロセスが先行していた(Bowers et al., 1990; Zander et al., 2015)
- ◇ 測定手法の制限のためである可能性がある

WRの分析では3or4回目の結果のみ、70%以上の試行が除外されている

⇒先行研究(Metcalf & Wiebe, 1987)では最大40回のWRの評価が行われており、データポイントを増やす必要があった

【補足：仮説と逆の結果が出た点に関しては、WRのデータ測定法が不適だったからという結論にしたようです。】

➤ 確信度評価

- ✓ 洞察試行では、自分の解が正しいという確信が全体的に強くなり、全体的に正確な回答が得られた。
- ✓ しかし、洞察においては、実際には確信度は正答率の予測にはあまり寄与しなかった。
- ✓ 非洞察試行で確信度と正答率の間に比較的強い対応が見られたことと、洞察試行では確信度が偶然のレベル以上の正答率を予測しなかった

⇒非洞察試行では参加者が問題解決に関するより多くのメタ認知的認識を持っていた

➤ 回答の安定性

参加者は、非洞察試行よりも、洞察試行の方が不正確な解答を保持していた

⇒非洞察試行の方が解答を調整しやすい

- ◆ 解答に至る前のメタ知識へのアクセス(WR)が洞察と非洞察で異なるという仮説は支持されなかったが、解答に到達した後に発生する直感的な感情(確信度)と意思決定戦略(多肢選択課題の回答)が異なるという結果は、2種類の問題解決が質的に異なるプロセスを反映していることを示す。

【無意識の知識を反映する洞察】

- ✓ 確信度と回答の正確さの間の対応関係は、行動が意識的な知識によって影響を受けていることを示しているのに対し、このような対応関係の欠如は通常、無意識的な知識利用を示唆する (Dienes and Berry, 1997)。
 - ◇ 洞察においてこの対応関係の欠如がみられたことから、洞察では関連する知識の意識的な利用が低い可能性を示唆
 - ◇ 非洞察回答では、確信度が回答の正確さを予測していたことも、メタ認知的に知識を利用を支持
- ◆ 洞察による問題解決における自動／無意識の認知プロセスを強調する理論（例えば、Ohlsson, 1992; Smith and Kounios, 1996; Knöblich et al., 1999; Topolinski and Reber, 2010）とより一致？

*WR の結果による裏付けもあればより強固なものだったが、今回は制限が多くそれは達成されなかった