

Salvi, C., Bricolo, E., Kounios, J., Bowden, E., & Beeman, M. (2016). Insight solutions are correct more often than analytic solutions. *Thinking & reasoning*, 22(4), 443-460.

Introduction

- 研究者は、アインシュタインの言う「大いなる推論的飛躍(“a great speculative leap”)」を、意識に解が突然現れるような”洞察(insight)”として一般化した。
 - ✓ 洞察は、解決に至るプロセスは無意識的であり、解は事実として意識的に再構築される (Metcalf & Wiebe, 1987; Smith & Kounios, 1996).
- もう 1 つの解決方略は“分析(analysis)”であり、解を見つけるための問題空間の意識的、慎重な探索を指す((Ericsson & Simon, 1993; Metcalfe & Wiebe, 1987; Newell & Simon, 1972).
 - ✓ 洞察解が all or none で解をもたらすのに対して、分析解はほとんど意識的に進行するためその途中結果を利用することが出来る (Kounios, Osman, & Meyer, 1987; Meyer, Irwin, Osman, & Kounios, 1988; Sergent & Dehaene, 2004; Smith & Kounios, 1996)..
- 洞察解は一度達成されると、その解が正しいものであるという強い感覚を持つ。
 - ✓ 実際に、偉大な発見の歴史は、突然現れた解が成功したエピソードを数多く残している。
 - ✓ ただし、このような信念は実験的にテストされたことはなく、肯定的な事例ばかりを報告し、機能しなかった洞察を無視している可能性がある
- ☆ 本研究は、彼らの洞察ともなう高い確信の正当性について検証する
- 洞察による解決策が分析的回答より正確だという仮説を疑う理由はいくつかある。
 - ① 分析処理は意識的に利用可能であり、段階的に進行し、進行状況を監視し、必要に応じて調整を行うことができるため、分析処理はより正確な解決につながると合理的に仮定できる
 - ⇨ 洞察解は意識の閾値以下で生成されるため、解が認識されるより前に、その処理を監視及び調整することはできない
 - ② 洞察における確信度は、洞察の経験に伴う確信度ではなく、洞察的に問題を解決する解決策を後づける的に洞察であるとラベル付けしているだけかもしれない
 - ✓ Metcalfe(1986)は、参加者が解に近づいたと感じる度合い(「暖かさの感覚(warmth)」)の評価が、洞察では解決前の最後の 10 秒間まで暖かさの評価が増加しなかったのに対し、分析の解決中の暖かさの評価はより漸進的に増加することを示した。
 - ✓ そして、Metcalf(1986)は、暖かさの感覚が突然増加した時のほうが、漸進的に増加した時に比べて正しいことを示した
- ☆ このことは、洞察・分析の差異が、解決策がもたらされるよりも前から生じていることを示す
- その他の、洞察と分析が異なるプロセスに基づくという知見
 - ✓ 神経イメージ研究では、洞察と分析の間の解決中および解決前の脳活動には異なるパターンがある (Jung-Beeman et al., 2004; Kounios et al., 2006; Subramaniam, Kounios, Parrish, & Jung-Beeman, 2009)

- ✓ 問題を見る前でさえ、その人の脳活動のパターンが、その問題が分析的に解決されるのか洞察的に解決されるのかを予測する (Kounios et al., 2008; Subramaniam et al., 2009)
 - ✓ 洞察と分析的解決の間で、瞬きや眼球運動に差異がある (Salvi, Bricolo, Franconeri, Kounios, & Beeman, 2015)
- 洞察と分析の間の認知方略の差異による解決中のエラーのパターン (Kounios et al., 2008)
- 洞察：commission エラー(誤った回答)ではなく omission エラー(タイムアウト)を犯す傾向がある
 分析：omission エラーではなく commission エラーをより犯す
 …制限時間が近くなったときに、洞察的な解決者は問題解決の途中の段階をモニターできないのでタイムアウトするが、分析的な解決者は制限時間直前に検証していた仮説を解として提案できるため誤答が多くなる
- 以上の知見は、洞察と分析の間の差異が、解決後のラベル付けの結果ではないことを示す。
- そこで本研究では、言語(CRA の問題)、視覚(物体の認識, or 「visual aha」問題)、または言語と口頭の混合問題(アナグラムとリバスパズル)を使用した 4 つの実験で、洞察と分析の間の相対的な正確さを比較した
- 基本的な実験手順は、以下の通り
1. 参加者は問題が提示されてから 15 秒(実験 2 では 16 秒)で、ボタンを押して回答することを求められた。
 2. 正解のフィードバックを受けることなく、回答した単語が洞察と分析のどちらによって行われたか回答させた
- 結果の予測
- (1) 洞察よりも分析的な解決の方が参加者の正答率が高い
 - (2) 分析よりも洞察的な解決の方が正確である
 - (3) 問題解決スタイルとその精度の間に関連はない

実験 1

参加者

- ノースウェスタン大学からの 38 人の学部生(ageM = 20.12, SD = 3.04, female = 22)が参加しました。すべての参加者は右利き。

刺激

- 120 の CRA 課題(Bowden & Jung-Beeman, 2003b; Mednick, 1962, 1968)のそれぞれについて、3 つの刺激語(crab, pine and sauce)が同時に提示された。
- 参加者はこれらの三つの単語に共通する単語(crab apple, pineapple, applesauce)を想起するように求められた。
- CRA 課題は洞察もしくは分析の解決方略によって解決される

- 自己報告による分類の妥当性は神経イメージ研究の知見からもその信頼性が実証されている(e.g., Bowden & Jung-Beeman, 2007; Jung-Beeman et al., 2004; Kounios & Beeman, 2014; Subramaniam et al., 2009)

手続き

- 実験前に,参加者には,3つのCRA問題と,洞察を分析的問題解決と区別する方法に関する指示が与えられた(Jung-Beeman et al., 2004; Kounios et al., 2006)
 - ✓ 洞察による課題の解決は,答えが突然頭に浮かび,驚きをとめない,解決策がどのように得られたのかを示すのが困難であると説明された
 - ✓ 分析による課題の解決は,ある単語の合成語を生成し,別の単語でテストするなどの戦略を使用して,解決策がどのように得られたかを述べるができるような,徐々に頭に答えが浮かぶものとして説明された.
 - ✓ さらに問題の難易度はさまざまであるため,すべてが解決できるわけではないと教示した.
 - ✓ 洞察・分析の報告について,正解も不正解もないことを伝え,および解決策のいずれも他のものよりも優れているとか,悪いとかがないことを伝えた
- 問題は,0.5秒の注視点の後にFig1Aのように提示された.
- 参加者は3つの単語を読み,解が思いついたらすぐにボタンを押すように求められた
- 解に対するフィードバックは与えられなかった
- 各回答の後に,参加者は洞察か分析かを判断し回答した.

結果

- 全体の時間枠の分析
 - ✓ 全体回答率 47.7%(回答数の平均 $M = 60.2$, $SD = 16.2$)
 - ✓ 洞察正答率 93.7%(回答数の平均, $M = 29.3$, $SD = 11.4$).
 - ✓ 分析正答率 78.3%(回答数の平均, $M = 19.1$, $SD = 11.2$).
 - ✓ 洞察正答率 > 分析的な正答率($t(37) = 5.66$, $d = .99$, 95%CI[.09, .21], $p < .001$)
 - ✓ 洞察におけるエラー < 分析におけるエラー($t(37) = -5.71$, $d = -1.02$, 95%CI [-.22,-.10], $p < .001$).
- 2~10sの時間枠の分析
 - ~2s : すぐ思いつく解を洞察として衝動的に報告する可能性がある
 - 10~15s : タイムアウト回避のために,分析的な回答がされやすい
 - ✓ 最後の5秒間(34.1% errors, mean n of errors, $M = 5.1$, $SD = 5.2$)の方が,それ以前の10秒間よりもエラー率が高かった(10.5% errors, mean n of errors, $M = 5.1$, $SD = 4.6$, $t(37) = -4.41$, $d = -.70$, 95% CI [-.20, -.07], $p < .001$).
- 洞察と分析間の比較を中心に限定することにより,これらの要因による正答率への影響を排除することが出来る
 - ✓ 全体に対するこの時間枠の回答率 77.8%(回答数の平均, $M = 45.2$, $SD = 11.6$)
 - ✓ 洞察正答率 94.4% (回答数の平均, $M = 25.4$, $SD = 10.3$).

- ✓ 分析正答率 83.2% (回答数の平均, $M = 14.5$, $SD = 10.1$).
- ✓ 洞察正答率 > 分析正答率 ($t(37) = 3.55$, $d = .74$, 95%CI [.04, .17], $p < .005$),
- ✓ 洞察におけるエラー < 分析におけるエラー ($t(33) = -3.61$, $d = -.81$, 95%CI [-.19, -.05], $p < .001$) (Figure 2)

実験1 考察

- 全時間と狭い時間枠の両方の結果で、分析よりも洞察のほうが正答率が高かった。
- この発見の一貫性は、残る3つの研究で検証していく
- 実験1では、参加者の報告により「分析」と「洞察」を分類させたが、実験2では「不明」を追加した。

実験2

方法

参加者

- 51人の学部生 ($M = 20.5$ 年, $SD = 2.7$, 女性25名)。参加者の脳波(EEG)を測定した(EEGデータはほかの論文において発表)
- すべての参加者は、右利き、アメリカ英語のネイティブスピーカー。

刺激

- 刺激は、コンピュータープログラム(Vincent, Goldberg & Titone, 2006)によって生成された180のアナグラム(109文字の4文字アナグラムと71文字の5文字アナグラム)
- すべてのアナグラムには1つの解しかないことが確認されていた。
- 課題の50%は実際にある単語(OWLS⇒SLOW)…単語刺激
- 残りの50%は実際にはない単語(LAGO⇒GOAL)…非単語刺激
- 平均バイグラム合計: 5954.91 ($SD = 2555.31$)
- 平均単語頻度(Francis & Kucera, 1982): 100万あたり 54.75 ($SD = 93.79$)
- 洞察と分析的解決を区別する自己報告は、EEGを使用した研究(Kounios et al., 2008)によって信頼性を担保されたものを使用。
- 刺激は、白い背景に黒い文字で提示(Fig1 B)

手続き

1. 実験前に練習ブロックを行った。
2. 教示は、実験1と同様でした。
3. 各試行は、0.5秒間の注視点が表示された後、画面の中央にアナグラムが表示された。
4. 参加者には、できるだけ早くマウスボタンを押すことで回答することが求められた。制限時間は16秒だった。
5. 回答の0.5秒後、画面メッセージが表示され、解決語を口頭で報告するように求められた。この時、フィードバックはなかった。
6. 回答後、自身の回答が洞察か分析か、不明かを分類した。

結果

- 全体時間枠における分析
 - ✓ 全体正答率 73% (回答数の平均, $M = 130$, $SD = 23.2$).
 - ✓ 洞察正答率 97.6% (回答数の平均 $M = 72.4$, $SD = 30.2$).
 - ✓ 分析正答率 91.9% (回答数の平均, $M = 51.9$, $SD = 30.1$).
 - ✓ 洞察正答率 > 分析正答率 ($t(50) = 2.78$, $d = 1.50$, 95% CI [.01, .09], $p < .01$)
 - ✓ 洞察におけるエラー < 分析におけるエラー ($t(50) = -2.79$, $d = .60$, 95% CI [-.1, -.06], $p < .01$).
 - ✓ 不明分類率 2.2% (回答数の平均, $M = 2.8$, $SD = 6.2$)

- 2~10s における分析(narrow time)
 - ✓ 実験1の時と同様, 最後の5s(11.1%; エラー数, $M = 2.5$, $SD = 1.6$)では, それ以前の11sにおける正答率よりも低くなっていた(4.3%; エラー数, $M = 3.8$, $SD = 4.1$; $t(45) = -4.3$ $d = -.76$, 95% CI [-.12, -.04], $p < .001$)
 - ✓ また, 最後の5sの回答は洞察(2.9%, $M = 0.6$, $SD = 1.1$)よりも分析(8.3%, $M = 1.9$, $SD = 1.6$)としてラベル付けされることが多かった($t(27) = -3.37$, $d = -1.08$, 95% CI [-.44, -.12], $p < .001$)
 - ✓ 全体に対するこの時間枠の回答率 68.7% (回答数の平均, $M = 89.2$, $SD = 16.3$).
 - ✓ 洞察正答率 97.6% (回答数の平均, $M = 48$, $SD = 5$),
 - ✓ 分析正答率 91.8% (回答数の平均, $M = 37.4$, $SD = 15.1$).
 - ✓ 洞察正答率 > 分析正答率 ($t(50) = 2.64$, $d = .51$, 95% CI [.01, .10], $p < .05$)
 - ✓ 洞察におけるエラー < 分析におけるエラー ($t(46) = -2.66$, $d = .53$, 95% CI [-.01, -2.66], $p < .05$).

実験2の考察

- 実験1の結果が再現された.
 - ✓ どちらの時間枠においても, 洞察とラベル付けした時, 分析とラベル付けした時に比べて, より正確な回答をしていた.
 - ✓ この知見は, どちらにも分類できない「不明」を入れたうえでも示された
- 実験3では, さらに課題の種類を変えたうえで, 異なる言語と文化においてもこの効果が再現されるかどうかを検討した.

実験3

方法

参加者

- Milano-Bicocca大学の学部生110人($M = 21.2$, $SD = 4.8$, 女性81名)が, 実験に参加した. すべての参加者は右利きでイタリア語のネイティブスピーカーでした.

刺激

- イタリアの rebus パズル(Salvi, Costantini, Bricolo, Perugini, & Beeman, 2015; see Figure 1(C))を

88 題使用した

- ✓ 各 rebus パズルを解決するためには、参加者は発話と視覚の手がかりを組み合わせ、次のような一般的なフレーズを作成する必要があった。
- ✓ Figure2 の c は LUNA が小さくなっていることから「luna calante(月の満ち欠け)」と答える必要があった
- ✓ Rebus パズルは洞察を引き起こすことが知られており、(MacGregor & Cunningham, 2008). CRA(実験 1)やアナグラム(実験 2)のように、洞察または分析的な回答をが可能(Salvi, Costantini, et al., 2015).

手続き

- 実験 1, 2 と同様の手順で行った

結果

- 全体の時間枠の分析
 - ✓ 全体の回答率 64.6%(回答数の平均, $M = 20.7, SD = 4.3$).
 - ✓ 洞察正答率 78.5%(回答数の平均, $M = 10.4, SD = 5$)
 - ✓ 分析正答率 63.2%(回答数の平均, $M = 5.3, SD = 4.7$)
 - ✓ 洞察正答率 > 分析正答率($t(109) = 6.08, d = .66, 95\% CI [.13, .27], p < .001$)
 - ✓ 洞察におけるエラー < 分析におけるエラー($t(109) = -3.71, d = -.72, 95\% CI [-.05, -3.7], p < .001$)
- 2~10 秒の枠内での分析
 - ✓ 実験 1, 2 と同様に最後の 5s と比較して、この 10s では正答率が高かった(エラー数 28.1%, $M = 3.6, SD = 18.3; t(108) = -4.82, d = -.61, 95\% CI [-.26, -.11], p < .001$).
 - ✓ 最後の 5s では、より多くのエラーが洞察(19.4%)よりも分析(39.2%)とラベル付けされた($t(109) = -4.82, d = -1.63, 95\% CI [-.5, -.18], p < .001$)
 - ✓ 2~10 秒の回答数(回答の 61.2%)
 - ✓ 洞察正答率 75.1%(回答数の平均, $M = 5.7, SD = 3.2$).
 - ✓ 分析正答率 64.4%(回答数の平均, $M = 3.4, SD = 2.9$).
 - ✓ 洞察正答率 > 分析正答率($t(109) = 4.19, d = .39, 95\% CI [.25, 4.1], p < .01$)
 - ✓ 洞察におけるエラー < 分析におけるエラー($t(103) = -2.10, d = -.39, 95\% CI [-.16, -.004], p < .05$)

実験 3 の議論

- 実験 3 も実験 1, 2 の結果を再現した
 - ✓ 全体の分析と、2~10s の分析の両方で、洞察とラベル付けされる解のほうが正確であった
 - ✓ 実験 3 の結果は、この傾向が問題の種類と、他の言語における頑健性を持つことを示す
 - ✓ 実験 4 では、より言語に依存しない視覚的な課題を用いた

実験 4

方法

参加者

- Milano-Bicocca 大学のイタリア語を母国語とする学生 27 名(M = 22.3 歳,SD = 1.9,女性 21 名)

刺激

- 刺激は, Snodgrass & Vanderwart(1980)におけるアニメーションおよび無生物オブジェクトの断片的な線画だった(Figure1(D))
 - ✓ 参加者は画像のセグメンテーションレベル 2 の画像を判別した(非常に少ない情報)
 - ✓ 画像セグメンテーションレベル…黒いピクセルを含むセグメントがランダムに累積的に削除されたもの(Snodgrass&Corwin,1988, Figure1(D)はレベル 2).
 - ✓ 洞察における近くの役割は古くから注目されており (Porter,1954, Rubin,Nakayama, & Shapley,1997,2002), 画像識別課題は洞察の問題解決を研究するための有効なタスクであることが実証されている(例: Ludmer,Dudai, & Rubin,2011)

手続き

- これまでの実験と同様の手続きをとった.

結果

➤ 全体時間枠での分析

- ✓ 全体回答率 71.8%(回答数の平均,M = 35.9,SD= 6.5)
- ✓ 洞察正答率 78.4%(回答数の平均,M = 16.6,SD = 6.7)
- ✓ 分析正答率 41.5%(回答数の平均,M = 6.7,SD = 4.7)
- ✓ 洞察正答率 > 分析正答率($t(26) = 7.47, d = 1.85, 95\%CI [.26, .46], p < .001$)
- ✓ 洞察におけるエラー < 分析におけるエラー($t(26) = -7.47, d = -1.85, 95\%CI [-.46, -.26], p < .001$)

➤ 2~10s の時間枠の分析

- ✓ 洞察正答率 68.6%(回答数の平均, M = 7.2,SD = 4.4)
- ✓ 分析正答率 43.6%(回答数の平均, M = 5.3,SD = 4.2)
- ✓ 洞察におけるエラー 31.4%(回答数の平均 M = 3.2,SD = 2.7)
- ✓ 分析におけるエラー 56.4%(回答数の平均,M = 6.3,SD = 3.56)
- ✓ 洞察正答率 > 分析正答率($t(25) = 4.36, d = 1.24, 95\%CI [.13, .36], p < .01$)
- ✓ 洞察におけるエラー < 分析におけるエラー($t(26) = -4.36, d = -1.08, 95\%CI [-.36, -.13], p < .001$).

実験 4 の考察

- 実験 4 も, これまでの結果を再現した
 - ✓ 全時間枠と 2~10s の時間枠の両方で,洞察とラベル付けされた回答のほうが, 分析とラベル付けされた回答よりも正確だった
 - ✓ 実験 4 の結果は, 視覚的な洞察課題においてもこの傾向が頑健であることを示す.

総合考察

- ▶ 少なくとも実験室で使用される問題において、洞察的な回答は分析的な回答に比べて正確だった
 - ✓ 洞察によってアナグラム課題を解決する場合、正しい解決を行うか、タイムアウトになる傾向があり、分析によって解決する場合、より誤答が多かったがタイムアウトすることはほとんどなかった。
 - ✓ この発見は、洞察力の解決は全か無かプロセスの結果であり、分析的な解決は漸進的な (incremental) プロセスの結果であるという知見と一致する (Smith & Kounios, 1996)。
 - ✓ 漸進的な解法で解いている場合には、参加者は制限時間の直前で、推論中の情報を利用して回答を行うことが出来るためタイムアウトは少なくなる
 - ✓ ⇨全か無かのプロセスである洞察の場合は、そのような中間的な結果を利用できないためタイムアウトが多くなる
- ☆ 最後の 5 s では分析とラベル付けされた誤った回答が非常に多かったという本研究の結果によって支持される。
- ▶ 洞察とラベル付けされる際に中間状態を報告できないのは、解が見つかるまでは、洞察は潜在的な連合の活性化により問題解決が進行するため、突然に解が認識されるためであると考えられる。
例) pine / crab / sauce における解決語 apple は、各単語と連合がとても強いわけではないため、その単語が発見されないうちは、意識上に上らないが、発見とともに 3 つのすべてとの連合が認識される
- ☆ 洞察的解決は中間状態をもたらさない全か無かのプロセスをとり、分析的解決は中間状態を利用可能な漸進的プロセスを取るために正答率に差異がみられる
- ☆ この傾向は、様々なタイプの課題において一貫して見られる

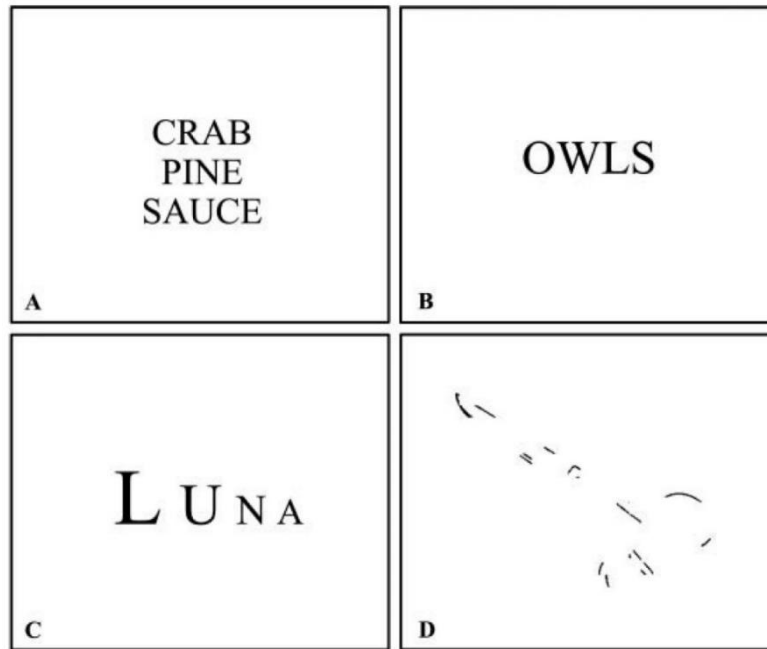


Figure 1. Each panel represents an example of the items used in the experiments. Specifically: 1(A) represents an example of CRAs, the solution is “apple”; 1(B) represents an example of anagrams, the solution is “SLOW”; 1(C) represents an example of rebus puzzle, the solution is luna calante “decreasing/waning moon” and 1(D) represents an example of visual aha problems, the solution is “scissors”.

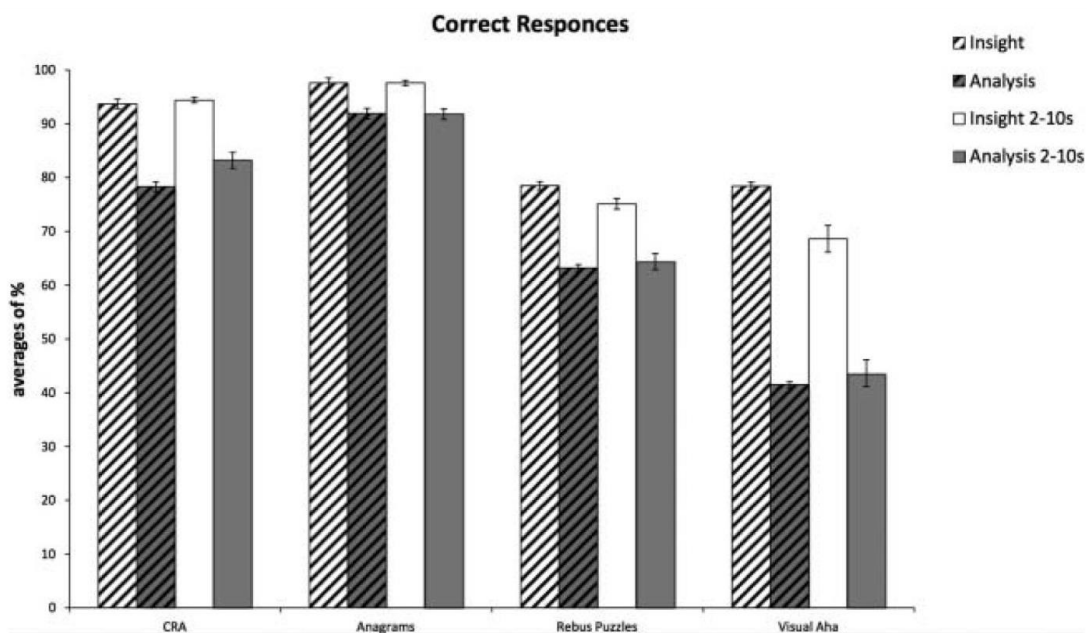


Figure 2. The per cent of problems correctly solved compared by type of problems (CRA, anagrams, rebus puzzle and visual aha), solution style (insight or analysis) and across two time windows (the whole problem or excluding the first 2 and the last 5 s). For each type of problem, significantly more correct problems were produced via insight compared to analysis. Errors bars indicate 95% confidence intervals. Note as mentioned in the text, 16 s were given to participants to solve the anagrams task; therefore, the second time-window considered goes from 2 to 11 s.