

The Role of Explanation in Discovery and Generalization: Evidence from Category Learning

Joseph J. Williams, Tania Lombrozo

Cognitive Science, 2010, 34, 776–806

1. Introduction

- ▶ 説明は学習を促進する (Roscoe & Chi, 2007, 2008)
- ▶ self-explanation effect (e.g., Chi, Bassok, Lewis, Reimann, & Glaser, 1989)
 - 説明相手がいなくても
- ▶ explanation-based learning (DeJong & Mooney, 1986; Mitchell, Keller, & Kedar-Cabelli, 1986)
 - 転移や一般化に大きな効果を持つ
- ▶ なぜ、説明することは学習、特に深い学習を助けるのか？
 - メタ認知、構造化、新しい知識の統合、メンタルモデルの繰り返し
 - (Chi et al., 1994; Chi, 2000; Siegler, 2002; Crowley & Siegler, 1999; Rittle-Johnson, 2006)

1.1. Exploring the role of explanation in generalization

- ▶ 説明の効果の **subsumptive constraints** 説明を調べる
 - 説明することは広い一般的な規則性の発見を促進するような制約を働かせる
- ▶ **pattern subsumption** 理論
 - 説明の属性を定義することは説明されるものがどのような一般的なパターンの事例となるかを示すこと
- ▶ 同じ一般的なパターンによる異なる観測の説明は **unification** を強調する
- ▶ 説明 = **subsumptive** 理論を満足させるべき = 一般的なパターンを見つける
 - 観測を統一的な規則によって解釈しようとする
 - 一般性の顕在的な認識
 - 転移の促進

◇ キーポイント

- ▶ 説明の構造的な属性を満たすことは学習者に統合的な規則性を発見させる制約を働かせるので、説明することは一般化を促進する

- ◇ 仮説のテストのために、カテゴリー学習課題を利用する
 - ▶ 理論的、実践的に既に説明とカテゴリー構造の間に重要な関係があるという証拠がある (e.g., Patalano, Chin-Parker, & Ross, 2006)
 - ▶ よく統制された材料をよく理解された課題で利用できる

2. Overview of experiments

- ◇ 主な仮説
 - ▶ カテゴリーのメンバーの説明を生成しようとすることは学習者に観察を一般的な共通のパターンで解釈することを強いる
 - 分かりにくい規則性の発見が促進される
 - ▶ 75%ルールと 100%ルール
 - 100%ルールのほうが発見が難しいが、カテゴリーのメンバーを最も統合的に説明できる
- ◇ 予測
 - ▶ 説明することは学習者に 100%ルールを発見させる

3. Experiment 1

- ▶ 説明条件と描写条件の比較

3.1. Methods

3.1.1. Participants

- ▶ 150名の学部生

3.1.2. Materials

3.1.2.1. Study items

- ▶ 2つのカテゴリー
 - GLORP と DRENT が各 4 体ずつ
- ▶ 4つの特徴
 - 左の色 (青、緑、赤、黄色)
 - 右の色 (茶色、シアン、グレイ、ピンク)
 - 体の形 (四角、丸)
 - 足の形 (8種類)
- ▶ 3つの分け方

- アイテムの類似性
- 75%ルール: 体の形で見分ける
- 100%ルール: glorp は足が尖っている

3.1.2.2. Test items

- ▶ 3種類 of アイテム (Figure 2)

3.1.2.3. Transfer items

- ▶ 完全に新しいアイテム

3.1.2.4. Memory items

- ▶ 23 のロボット
 - 8つは以前と同じ
- ▶ 15 はルアー

3.1.3. Procedure

3.1.3.1. Introduction phase

- ▶ 設定
 - 惑星 Zarn からきたロボット Glorp と Drent

3.1.3.2. Study phase

- ▶ 8つの study items がラベルとともに表示される
 - 各ロボットにつき 50 秒
- ▶ 説明条件: なぜロボットがそのタイプなのかを説明する
- ▶ 描写条件: ロボットのタイプの特徴を描写する

3.1.3.3. Test and transfer phases

- ▶ test items がランダムに表示される
- ▶ 続いて transfer item が表示される
- ▶ それぞれを各ロボットに分類する
- ▶

3.1.3.4. Memory phase

- ▶ memory item が表示される

- ▶ study phase で見たロボットか判断する

3.1.3.5. Explicit report

- ▶ どうやって見分けたかを聞かれる

3.2. Results

3.2.1. Basis for categorization

- ▶ 言語報告を4つのカテゴリーに分類 (Table 2)
 - 100%ルール、75%ルール、item similarity、other
- ▶ ある分け方とそれ以外の2つに分け分析
 - 説明条件の参加者は描写条件の参加者より100%ルールに分類されやすい ($\chi^2(1) = 15.89, p < .001$)
 - 描写条件の参加者は説明条件の参加者より75%ルールに分類されやすい ($\chi^2(1) = 19.56, p < .001$)
 - その他は有意差なし

3.2.2. Categorization of test and transfer items

- ▶ どのルールに基づいて test、transfer items を分類したか (Figure 3)
 - 100%ルールに基づいて分類すれば正解
- ▶ 正しさに対して2(課題: 説明 vs. 描写) × 2(分類測定法: テスト vs. 転移) の分散分析
 - 課題の主効果: 説明 > 描写 ($F(1,148) = 16.10, p < .001$)
 - 分類法測定の主効果: テスト > 転移 ($F(1,148) = 13.46, p < .001$)
 - 説明条件の参加者のほうが転移アイテムをより正しく分類した ($t(148) = 2.91, p < .01$)
- ▶ test items のタイプごとに分析
 - 100% rule probes を100%ルールで分類
 - ・ 説明 > 描写 ($t(148) = 4.41, p < .001$)
 - 75% rule probes を75%ルールで分類
 - ・ 描写 > 説明 ($t(148) = 3.77, p < .001$)
 - item similarity probes には差はない ($t(148) = 1.37, p = .17$)

- ▶ ノンパラメトリック分析
 - 転移アイテムの分け方により参加者を分類
 - ・ 8個中7個以上が正しい: 100%ルールを使用
 - 説明することは100%ルールの発見を促進する ($\chi^2(1) = 10.37, p < .01$)

3.2.3. Memory for study items

- ▶ 説明活動は他と異なるアイテムへ注意を向けさせる
- ▶ 75%ルールにあうアイテム (consist) とあわないアイテム (anomalies) に分けて分析
- ▶ 感度 d' 指標で表現 (Figure 4)
 - consistent: 説明 < 描写 ($t(148) = 2.24, p < .05$)
 - anomalous: 説明 = 描写 ($t(148) = 0.82, p = .41$, Explain: 1.09, Describe: 0.81)
- ▶ ルール報告の分類により分析 (Figure 5)
 - 100%ルール、75%ルール: 有意差なし ($t(30) = 0.88, p = .88, t(52) = 1.41, p = .16$)
 - other: 説明 < 描写 ($t(62) = 2.19, p < .05$)
 - ・ ルールの発見ではなく説明と描写の間の詳細のコーディングの効率の違い

3.2.4. Coded content of explanations and descriptions

- ▶ study phase における記述を分類 (Figure 6)
 - 各特徴に言及されているか (足の形、体の形、色)
 - どのように言及されているか
 - ・ abstract: より一般的な描写 (尖った/平ら、寒色/暖色)
 - ・ concrete: 実際の特徴 (三角/四角/L字、四角/丸)
- ▶ concrete と abstract それぞれで言及数に 2 (課題: 説明 vs. 描写) \times 3 (特徴: 足 vs. 体 vs. 色) の分散分析
 - abstract: 説明 > 描写 ($F(1,148) = 24.72, p < .001$)
 - concrete: 説明 < 描写 ($F(1,148) = 164.65, p < .001$)
- ▶ 説明 = abstract、描写 = concrete
 - 体の abstract ($t(148) = 0.82, p = .41$) をのぞきそれぞれの特徴で同じ結果 (all $ps < .05$)
- ▶ 足の形に言及したのは描写条件のほうが多かったが、説明条件の参加者のほうがより100%ルールを見つけた

3.3. Discussion

- ▶ 描写することに比べて、説明することのほうがわかりにくい規則性の発見を促進した
- ▶ 説明は **abstract** に分類される言及をより多く含んだ
- ▶ 描写は **concrete** に分類される言及をより多く含んだ
 - 単に特徴に注目するだけでは規則性の発見には不十分
- ▶ 描写は詳細のコード化により記憶テストのパフォーマンスが良かった
 - 発見と記憶の乖離
- ▶ 説明における **subsumption** と **unifying** の役割の証拠
 - 説明をすることによりカテゴリーのメンバーのより統一的な規則性を見つけようとした
- ▶ **self-explanation** 効果のカテゴリー学習への拡大
- ▶ **abstract** な描写 → 説明は素材をより多様で、曖昧な方法で表すことを促し、統一的な規則性の発見を促進する

4. Experiment 2

◇ 目的 1

- ▶ 説明が発見を促進することをより強く示す
 - 声に出して考える条件を統制群とする
 - 全ての参加者が後に、ロボットを分類する課題をすることを伝えられた

◇ 目的 2

- ▶ **anomalous** の役割を調べ、発見の過程をさらに調べること
- ▶ 可能性 1
 - 75%ルールにあわないロボットに気づけば 100%ルールを見つける
 - 説明は **anomalous** の発見を促進する
- ▶ 可能性 2
 - 説明に従事することは **anomalous** に直面するかに関わらず参加者に規則性を考えることを促す
- ▶ 可能性 3
 - 思考だけでは最初に 75%ルールを見つけることにより 100%ルールの発見が抑制される
- ▶ 支持する可能性
 - 説明と **anomalous** の観察の結合が、メンバー全体を包括する学習の制約により、

発見へ導く

4.1. Methods

4.1.1. Participants

- ▶ 140名の学部生

4.1.2. Materials

- ▶ 実験1と同じ
(study items と memory item を少し変更)

4.1.3. Procedure

4.1.3.1. Task instructions

- ▶ 後にロボットを思い出すテストと見分けるテストをすることを教示する

4.1.3.2. Prestudy exposure

- ▶ 8体のロボットを4秒ずつ表示 × 3

4.1.3.3. Study phase

- ▶ 2体のロボット(各グループから1体)を90秒ずつ表示
- ▶ consistent 条件: 75%ルールと一致
- ▶ anomalous 条件: 75%ルールと不一致
- ▶ 説明条件: 実験1と同様(発話)
- ▶ 発話思考条件: 思ったことを発話するように教示

4.1.3.4. Test, transfer, and memory

- ▶ 実験1と同様(2秒以内に反応)

4.1.3.5. Postexperiment questions about body shape

- ▶ いくつの Glorp (Drent) の体が四角かった(丸かった)か?

4.2. Results

4.2.1. Basis for categorization and categorization accuracy

- ▶ 言語報告に基づく分類 (Table 3)

- ▶ 説明することは、声に出して考えることより 100%ルールの発見を促進する
- ▶ 課題 × アイテムタイプ × 発見 の階層的対数関数分析
 - 課題と発見の交互作用 ($\chi^2(1) = 21.91, p < .001$)
 - 100%ルールの発見は 両説明条件においてどちらの思考発話条件より多い
($\chi^2(1) = 8.71, p < .01, \chi^2(1) = 8.71, p < .01$; and $\chi^2(1) = 13.30, p < .001, \chi^2(1) = 13.30, p < .001$)
- ▶ 分類の正しさ (Figure 7)
- ▶ 2(課題: 説明 vs. 思考発話) × 2(アイテムタイプ: 一致 vs. 不一致) × 2(分類測定法: テスト vs. 転移) の分散分析
 - 課題の主効果: 説明 > 思考発話 ($F(1,236) = 21.90, p < .001$)
 - アイテムタイプ的主効果 傾向 ($F(1,236) = 3.35, p = .07$)
 - アイテムタイプと課題の交互作用 傾向 ($F(1,236) = 3.35, p = .07$)
 - 分類測定法の主効果: テスト > 転移 ($F(1,236) = 14.38, p < .001$)
 - 分類測定法と課題の交互作用 ($F(1,236) = 4.71, p < .05$)
 - 分類測定法とアイテムタイプの交互作用 ($F(1,236) = 4.71, p < .05$)
- ▶ 説明条件の不一致条件が他と比較してより正しい分類をした
 - vs. 説明 & 一致 ($F(1,118) = 5.83, p < .05$)
 - vs. 思考発話条件 ($F(1,118) = 14.51, p < .001, F(1,118) = 12.68, p < .001$)

☆ ノンパラメトリック分析

- ▶ 転移アイテムの分け方により参加者を分類 (Table 4)
- ▶ 8 個中 7 個以上が正しい: 100%ルールを使用
- ▶ 課題 × アイテムタイプ × 発見 の階層的対数関数分析
 - 課題とアイテムタイプの交互作用 ($\chi^2(1) = 18.59, p < .001$)
 - アイテムタイプと発見の交互作用 ($\chi^2(1) = 3.91, p < .05$)
- ▶ 説明と不一致の組み合わせが発見を促進する
 - 説明条件におけるアイテムタイプの効果が有意傾向
 - 不一致 > 一致 ($\chi^2(1) = 2.76, p = .10$)
- ▶ 全ての指標において、説明することは発見を促進した
- ▶ 部分的に anomalous の効果が見られた

☆ 体の形をどれほど使用したか

- ▶ 2 次の交互作用 課題×アイテムタイプ×体の形の使用 ($\chi^2(1) = 4.35, p < .05$)
 - 説明条件において、一致条件で不一致条件よりも 75%ルールを使用した傾向がある ($\chi^2(1) = 3.68, p = .055$)
 - 思考発話条件では差がない ($\chi^2(1) = 0.95, p = .33$)

4.2.2. Memory

- ▶ 各判別アイテムにおいて d' を用いた 2 (課題: 説明 vs. 思考発話) × 2 (アイテムタイプ: 一致 vs. 不一致) の分散分析
 - アイテムタイプの主効果
 - ・ anomalous アイテムの判別で 一致 < 不一致 ($F(1,236) = 21.53, p < .001$)

4.2.3. Postexperiment questions about body shape

- ▶ 四角(丸)い体のロボットの数
 - 4 と回答したら anomalous に気づいていない
- ▶ ほとんどの参加者が anomalous に気づいていた

	一致	不一致
説明	74% (20/27)	92% (23/25)
思考発話	65% (17/26)	64% (16/25)

4.3. Discussion

- ▶ 説明することは思考を発話することより発見を促進する
 - テストがあることを知らされても変わらない
- ▶ anomalous を説明することは適切な学習と一般化により効果的
 - うまくいかない信念を棄却する
 - anomalous に注目したり、思考を発話するだけでは不十分である
- ▶ 記憶成績に差がない → 実験 1 の記憶の差は描写により促進された

5. Experiment 3

- ▶ 何も行わせない統制条件と比較
- ▶ 全ロボットを同時に提示

5.1. Methods

5.1.1. Participants

- ▶ 120名の学部生

5.1.2. Materials

- ▶ 実験 1,2 と同じ
 - それぞれに 1~8 の番号がふってある
 - テストと転移のアイテム数を減らす

5.1.3. Procedure

- ▶ 基本的に実験 1 と同様
- ▶ 8体のロボットが2分間表示される
 - 説明条件: カテゴリーの理由を説明
 - 自由学習条件: カテゴリーを教えられるだけで後は自由
- ▶ 質問
 - 一つの特徴でロボットを正しく見分けられる可能性は?
 - ・ 0, 25, 50, 75, 100%から選択
 - 二つのカテゴリーに見つけられる違いはあったと思うか?
 - ・ あるなら、その違いは何か?
 - 色や形の確認 (結果は省略)
 - 分類するのにどの特徴を使ったか?
 - 8つのロボットが表示されたときに、それぞれのロボットの分類理由を説明しようとしたか?

5.2. Results and discussion

5.2.1. Basis for categorization and categorization accuracy

- ▶ 言語報告に基づく分類 (Table 5)
- ▶ 説明することは 100%ルールの発見に効果かがある
 - 言語報告 ($\chi^2(1) = 4.09, p < .05$)
 - テスト、転移課題 ($F(1, 118) = 7.02, p < .01$)
- ▶ 説明することは 75%ルールの使用を減らす ($\chi^2(1) = 4.66, p < .05$)

5.2.2. Memory

- ▶ どちらのアイテムも有意差なし

- consist item ($t(118) = 1.60, p = .11$: Explain: 0.89, Free study: 0.35)
- anomalous item ($t(118) = 0.26, p = .80$: Explain: 0.74, Free study: 0.62)

5.2.3. Likelihood of underlying feature

- ▶ 一つの特徴でロボットを正しく見分けられる可能性は？
 - 有意差なし ($t(118) = 0.65, p = .52$: Explain: 37.9, Free Study: 42.1)
 - 発見ルールのカテゴリーごとに分析しても (all $ps > .12$)

5.2.4. Self-report of explaining

- ▶ 分類理由を説明しようとしたか？ (Table 6)
 - 条件間に有意差なし ($\chi^2(1) = 2.41, p = .30$)
- ▶ 発見(足の形を見つけたか) × 課題(説明 vs. 自由) × 報告(yes vs. not sure & no) の対数線形分析 (Table 7)
 - 課題と発見の交互作用 ($\chi^2(1) = 4.17, p < .05$)
 - 報告と発見の交互作用 ($\chi^2(1) = 13.96, p < .001$)
- ▶ 教示されてでも、自発的にでも説明をすることは規則の発見を促進する

6. General discussion

☆ 3つの実験結果より

- ▶ 参加者になぜアイテムがあるカテゴリーに属するのかを説明させることはより一般的なルールの発見を促進する
- ▶ **subsumptive constraints** 説明を支持
 - 良い説明とはどのような一般的なパターンの事例であることを示すこと
 - カテゴリーメンバーの説明をすることはパターンや規則性の発見を引き出す
 - 説明がより多くの観測を説明するほうが良い
 - より広い一般性の推論

☆ 実験 1、2 の結果は説明をすることが、抽象化と **anomaly** の処理の促進を示す

- ▶ 抽象化により、多くの事例に統一的な説明ができる
- ▶ **anomaly** を説明することは一致しない信念の棄却に導く

☆ 説明が役に立たない状況もある

- ▶ データが不十分なときや完璧なルールがないとき

- ▶ 今後の研究
 - バイアスや勘違いはこの観点からどのように説明されるか？
 - 説明が適切な学習を阻害する状況での検討

6.1. Alternative interpretations of the effects of explanation

- ▶ 実験 1,2,3 は互いに補完し合っている

- ◇ その他の代替説明
 - ▶ 説明をさせることが暗に参加者にカテゴリーのつながりを探ること（実験車の意図など）を伝えている
 - 実験 2,3 では後の分類テストを伝えているが説明の効果がある
 - 説明をさせなくても 75%ルールは見つけている
 - 説明をさせたことによる違いはより統一的で包括的ルールを見つけたこと

6.2. Relationship between explanation and other cognitive processes

- ◇ 他の認知プロセスとの関係

- ◇ 記憶探索のプロセスの深さ
 - ▶ 説明をした参加者は記憶という点では描写条件より劣っている
 - 記憶の深い処理とは異なる深い処理が必要

- ◇ 規則学習
 - ▶ 説明はカテゴリー学習におけるルールベースの方略を増加させる
 - ▶ ただし、説明はいろいろな制約を起こすためルールベース方略と同じではない

- ◇ 仮説生成の観点からの理解
 - ▶ 説明条件の参加者は仮説の生成、検証を行っていたのかもしれない
 - ▶ アイテムの比較を行っていたのかもしれない
 - ▶ 排他的ではなく相互補完的な提案である
 - 説明をすることは特定の制約（統一的なパターンで解釈する）を課す
 - そのような制約を満たすために、仮説生成、検証、アイテムの比較を行う
= 制約を満たすためにいろいろな認知プロセスを利用する

7. Future directions and conclusions

- ▶ どのような種類のパターンや規則性が説明的と判断され、規則性として発見されるか
- ▶ 説明することは因果的な規則性に特に有効か？
- ▶ 知識の獲得において、メカニズム的、機能的な説明はそれぞれ異なる役割を果たすか？
- ▶ 単純な説明を好む傾向は学習に影響するか？
- ▶ 機械学習への応用