

2014/12/1(月) ←11/26(水)分

輪講 担当：水野

## Tradeoff between Problem-solving and Learning Goals:

### Two Experiments for Demonstrating Assistance Dilemma

Kazuhisa Miwa, Hitoshi Terai, Ryuichi Nakaike, Cogsci2012

#### Abstract

学習支援システムは、参加者に様々なタイプのサポートを提供する。

仮説：高レベルサポートは参加者の問題解決目標への方向付けを活性化させるが、学習目標達成の優先度を低下させる。高い問題解決成績が得られるが、学習効果は低い。

素材にハノイの塔のパズルと自然演繹課題を用いた実験で、仮説は支持された。

#### Introduction

- 最近では、認知科学の理論に基づくインタラクティブな学習支援システムが開発されている
  - インタクションでは、学習支援システムと学習者のデザインや、学習者へのフィードバックが中心的な問題点である
- アシスタントジレンマ
  - 「支援をするのとなしないのとで、ベストな学習を達成するには学習環境のバランスはどうしたらよいか」(Koedinger and Alevan,2007)
  - 手厚い支援：成功への手がかりを提供し学習を向上させるが、生徒がよく考えることなく表面的な反応をする
  - 手厚くない支援：生徒に大きな苦勞をさせるが、効果的な学習につながる多くのエラーと干渉
- 本論文でのアシスタントジレンマの定義：問題解決を目標とするか学習を目標とするかを選択することのトレードオフ
  - 問題解決の目標を達成することは、学習支援システムで例題をできるだけ正確に速く解くことを意味する
  - しかし、主要な目標は例題を解くことではなく、例題を解くことから学ぶこと
  - 高いモチベーションの子供はただ問題を解くよりも、新しいものを理解したり、マスターするための能力を向上させる学習目標を設定する傾向にある。(Dweck,1996;Ames,1992)
  - 問題解決目標は学習支援システムのサポートによって達成されるが、学習目標は学習支援システムのサポートなしで達成される
- フィードバック
  - 参加者の目標設定は支援システムのフィードバック情報によって影響される
  - 直接的なフィードバックは次のステップで何が修正される必要があるかを教える。→具体的で問題解決に直接的に関係する
  - タイミング(即座 or 遅延)
    - ◇ タイミングの効果は課題の難しさと生徒のニーズや性格といったほかの要因と相互に作用して現れる(Mathan & Koedinger,2002)
    - ◇ 即座のフィードバック参加者が問題を解くときに何をすべきか決定する方向性を繰り返し与えるため

に、問題解決目標を設定することを促進するかもしれない

- 本研究について
  - 支援レベルを調整
  - 高レベルサポート＝直接的,即座のフィードバック
  - 仮説：高レベルサポートは参加者が問題解決目標を立てることを促進させ、学習目標を達成しようとする優先順位を低下させる。→問題解決の成績は低レベル支援条件よりも高いが、学習効果は低い

### Experiment1

課題：ハノイの塔(6 ディスク)(TOH)

実験システム

- 実験室で一人一台の PC。参加者はマウス操作で可能なディスクの動きを選択する
- プロダクションシステム上で実験を行う
- ベストな 1,5,9 手先の状態をヒントとして表示
- 支援レベル：ヒントで調整
  - 1 手先状態提示条件 > 5 手先状態提示条件 > 9 手先状態提示条件 > 支援なし条件

参加者と手順

- 71 人(17,19,17,18 人)
- 実験時間 90 分
- ハノイの塔を解くための戦略を学習するように教示され、学習セッションの後にポストテストでスキル獲得の度合いをテスト
- システムの使い方とルールの学習→ラーニングフェーズ(60 分・4 条件分岐)→ポストテスト(ヒントなし)

結果

- 問題解決成績(Figure2)
  - 用いた指標：ラーニングフェーズでの標準化ステップ数=(参加者が実際に従う必要があったステップ数)/(最小のステップ数)
    - ◇ 1.0 が最高の問題解決。1.0 より大きいとそれだけ無駄にステップ数を重ねたことになる
  - 最も低い支援レベルで最も悪かった
  - 支援あり 3 条件では有意な差は見られなかった。→ヒントが 9 手先状態提示条件でも十分すぎたのかもしれない

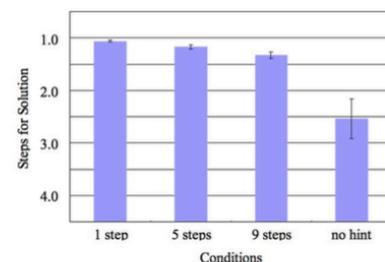
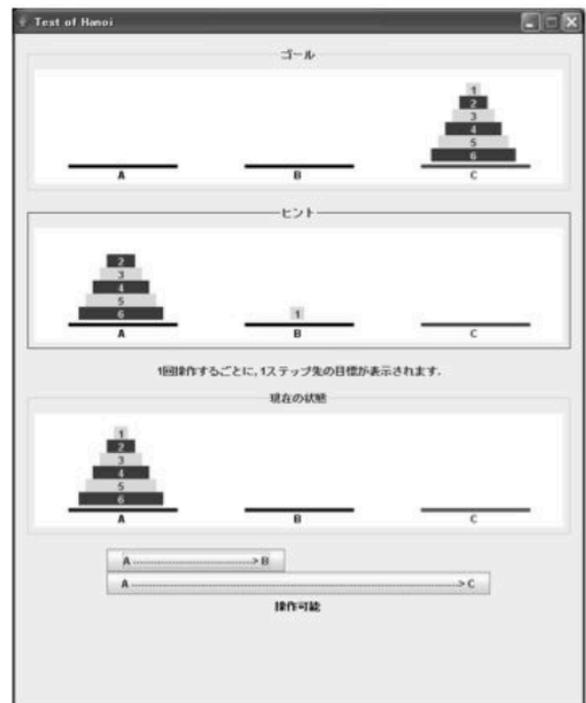


Figure 2: Average steps for solving problems in learning phase (TOH). The value of the vertical axis is reversed to represent higher values as high problem-solving performances, i.e., value 1.0 means the optimal problem solving. An ANOVA indicates that the main effect of Conditions factor was significant ( $F(3, 67) = 12.5, p < 0.01$ ). Fisher's LSD analysis shows significant differences between one step and no hints ( $p < 0.01$ ), five steps and no hints ( $p < 0.01$ ), and nine steps and no hints ( $p < 0.01$ ).

2014/12/1(月) ←11/26(水)分

輪講 担当：水野

- 思考時間(Figure3)
  - 用いた指標：ディスクの動きを決める平均時間
  - 学習目標よりも問題解決目標が優先されていることは、思考時間の減少を意味すると仮定
  - 1手先状態提示条件が5,9手先状態提示条件よりも短かった→仮説確認
  - しかしながら、ヒント無し条件では、時間が5,9手先状態提示条件よりも短い→仮説と矛盾する。

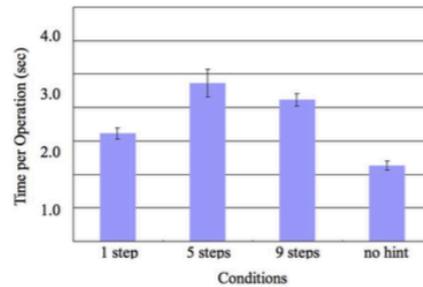


Figure 3: Average time for deciding disk movement (TOH). An ANOVA indicates that the main effect of Conditions factor was significant ( $F(3, 67) = 18.0, p < 0.01$ ). Fisher's LSD analysis shows significant differences between one and five steps ( $p < 0.01$ ), one and nine steps ( $p < 0.01$ ), five steps and no hints ( $p < 0.01$ ), and nine steps and no hints ( $p < 0.01$ ).

- 学習成績(Figure4)
  - 用いた指標：ポストテストでの標準化ステップ数
  - 低いレベルの支援ほど、学習効果は高い→仮説を確認
  - しかしながら、ヒント無し条件で、成績は5,9手先状態提示条件よりも悪い→仮説と矛盾→ヒント無し条件では難しすぎたのかもしれない

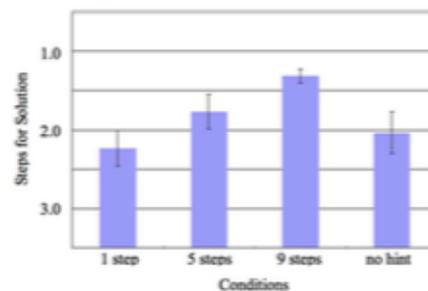


Figure 4: Average steps for solving problems in posttest (TOH). The value of the vertical axis is reversed to represent higher values as high problem-solving performances, i.e., value 1.0 means the optimal problem solving. An ANOVA indicates that the main effect of Conditions was significant ( $F(3, 67) = 3.25, p < 0.05$ ). Fisher's LSD analysis shows significant differences between one and nine steps ( $p < 0.01$ ) and nine steps and no hints ( $p < 0.05$ ).

## Experiment2

課題：自然演繹課題(ND)

- 証明問題
- 例： $\neg Q \rightarrow \neg P$  は  $P \rightarrow Q$  から導ける
- 基本的な9つの推論規則と5つの戦略でほとんどの問題が解ける

実験システム

- 学部生に自然演繹を教えるシステムとして開発された学習支援システム(Figure5)
- 参加者に推論規則と戦略を提供する
- リストから選ぶと、システムは自動的にそのルールを実行し、推論結果を表示する
- 規則と戦略の選択についてのヘルプ情報を与え、適用されるべきルールの候補を提供することで支援する

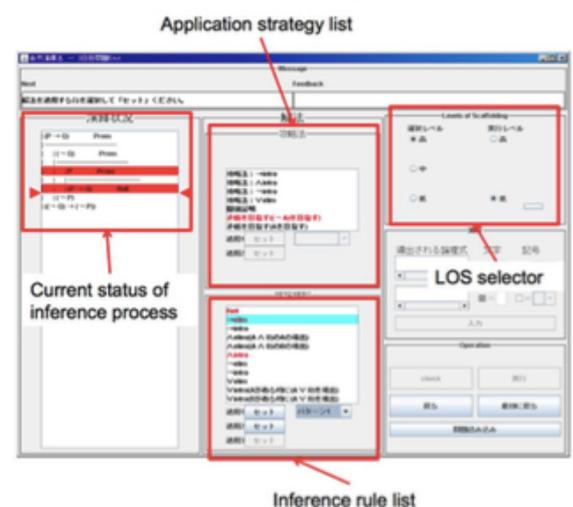


Figure 5: Example screen shot of ND tutoring system. The left side window shows a current status of inference processes where the red items are propositions to which the selected inference rule should be applied. The upper and lower center windows show a strategy list and an inference rule list where the red items are applicable candidates. The right upper window shows the LOS selector where the levels of support are controlled. In the current experiment, LOS was fixed based on the experimental manipulation.

2014/12/1(月) ←11/26(水)分

輪講 担当：水野

- 高レベル支援条件：推論規則と戦略の適用可能な候補を提供し、どの規則を適用するかを提案を表示。
- 低レベル支援条件：推論規則しか表示しない

参加者と手順

- 参加者：29人(高レベル13人,低レベル16人)
  - 手順：3週間にわたって行われた
    - 1週目：参加者はシステムの例としてシステムと推論の基本を学んだ。
    - 2週目(学習フェーズ前半)：参加者が4つの基本的な推論ルールを学んだ。全員が高レベル支援条件。
    - 3週目：支援レベルが操作。参加者は比較的複雑な問題(サブゴール設定が必要)を4問解いた(学習フェーズ後半)。
- 2つのポストテスト。ポストテスト1は学習フェーズで説いた問題2と同じで、ポストテスト2は初見の問題。

結果

- 問題解決成績(Figure6)
  - 用いた指標：学習フェーズ後半での平均時間
  - 問題3,4を解いている時に、低レベル支援より高レベル支援のほうが解く時間が短い→仮説を確認
- 思考時間(Figure7)
  - 用いた指標：学習フェーズ後半での推論規則を決める時間
  - 問題3,4では、低レベル支援よりも高レベル支援のほうが短い

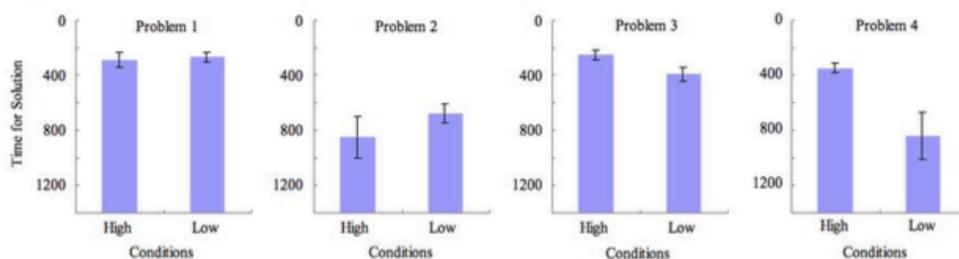


Figure 6: Average solution times for solving problems in learning phase (ND). The value of vertical axis is reversed to represent higher values as high problem-solving performances. T-tests show a marginal significant difference between high and low conditions in Problem 3 ( $t(22) = 1.92, p = 0.07$ ) and a significant difference in Problem 4 ( $t(15) = 2.70, p < 0.05$ ), but no differences in Problems 1 and 2 ( $t(25) < 1, n.s.$ ;  $t(24) = 1.14, n.s.$ ).

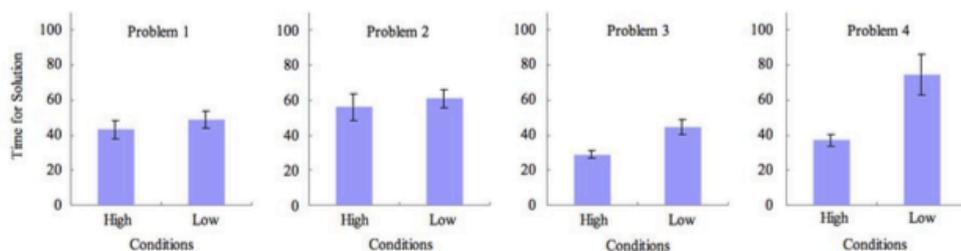


Figure 7: Average time for deciding inference rule (ND). T-tests show a marginal significant difference between high and low conditions in Problems 3 ( $t(22) = 2.55, p < 0.05$ ) and 4 ( $t(25) = 2.98, p < 0.01$ ), but no differences in Problems 1 and 2 ( $t(25) < 1, n.s.$ ;  $t(24) < 1, n.s.$ ).

- 学習成績(Figure8)
  - 用いた指標：ポストテストでの正解率
  - 初見の問題(Posttest2)で低レベル支援の成績が高い→仮説を確認

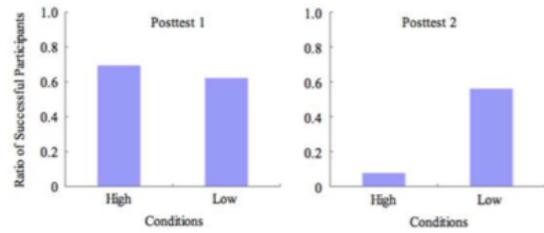


Figure 8: Ratio of successful participants (ND) . Chi square tests show a significant difference between high and low conditions in Posttest 2 ( $\chi^2(1) = 7.49, p < 0.01$ ), but no difference in Posttest ( $\chi^2(1) < 1, n.s.$ ).

### Discussion and Conclusion

- Koedinger et al.(2008) : 支援レベルの逆 U 字曲線 (Figure9)
  - 支援レベルが上がるにつれ、問題解決成績は向上するが、学習効果はある時点で最大になり、そこから減少する
- 超低レベル支援 or 支援なしは問題解決成績も学習効果も低下

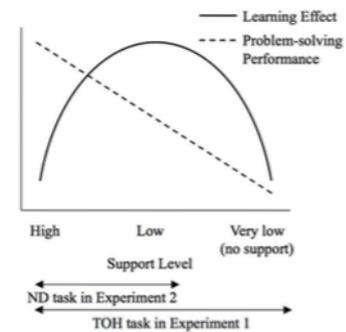


Figure 9: Conceptualized diagram of assistance dilemma and relation of problem-solving and learning goals.

- これは、たとえ参加者がゴールを設定しても、何をしたらよいかかわからず、多くの間違いをした結果、学習効果が少なくなってしまったと考えられる。
- 参加者がサポート無し学習で困難に直面したので、学習目標を達成することを諦めたのかもしれない。
- 現在の研究では、どちらの考察が正しいか決められない。
- アシスタントジレンマ=トレードオフの関係は多くの研究分野で見られる
  - Burns & Vollmeyer, 2002; Sweller, 1988
    - ◇ 参加者はある特定の目標を達成すると現象の裏にある理論や規則を考えない。
    - ◇ 参加者は、目標を達成するための空間における探索に集中しているので、仮説空間での探索を行わない傾向にある。結果として、その課題のルールや裏で動いているメカニズムを見つけられない。
    - ◇ この傾向は、目標の詳細さを操作することで調整できる。
  - 自動化システムの依存 Sheridan & Verplank, 1978
    - ◇ 原子力発電所の操作を行う実験において、自動化システムを継続的に使っていると、自動化システムが無効になった時に深刻な被害を起こす(Lin, Yenn, & Yang, 2010)
    - ◇ 実験システムでは自動化システムの不具合を見つけることは、不完全な自動化システムよりも安定している自動化システムで見つけることが難しい(Parasuraman, Molloy, & Singh, 1993; Molloy & Parasuraman, 1996)