

Exploration strategies, performance, and error consequences when learning a complex computer task

Van der Linden, D; Sonnentag, S; Frese, M; Behaviour & Information Technology, 2001, Vol.20, NO. 3, pp.189-198

1. Introduction

1.1. Problem solving strategies

- 複雑な問題解決と人間信頼性評価に関する研究は、問題解決の際に使用される方略に着目する。
- 先行研究により、体系的な探索 (Systematic exploration)、試行錯誤 (The trial-and-error strategy)、リジッド的な探索 (Rigid exploration)、情報検索における encapsulation は区別されている。
- 先行研究により、体系的な探索は、最善の方略であり、この方略を用いて問題を解決できる可能性が一番高い。
- 「体系的」というのは、タスクがどのように行われるべきかについての仮説を論理的方法でテストして、その結果に基づいてさらなる行動を計画するプロセスを指す。
- 試行錯誤方略は、効果が低いまたは効率が低いと想定される。
- この方略は、どのように行動するかに関する仮説がなく、次の行動が予測不可能あるいはランダムである。
- 試行錯誤方略の使用を導く原因はいくつかある。
- 先行知識がないのは原因の1つである。
- 先行知識がないと、どこで答えを探せるかわからないや計画を立てられない。
- もう1つ可能性のある原因は、タスクに対するコントロールがなくなり、ネガティブな感情が生じるからである。
- 複雑な問題を解決するプロセスで、エラーが起こる場合がよくある。
- エラーの発生は、自己効力感の低下やネガティブな感情の生じをもたらす可能性がある。
- 人は、自己制御 (self-regulation) により、ネガティブな感情を扱うことができる。
- 自己制御がうまくできない場合、多くの認知資源はネガティブな感情を処理するために消耗されてしまう。
- その結果、問題を解決する際に使える認知資源が少なく、計画の作成や結果の判断ができなくて、試行錯誤に至る。
- 問題解決のプロセスで、リジッド的な探索はよく観察される。
- リジッド的な探索の特徴は、アクションシーケンスの繰り返しである。
- ある仮説を確信して、たとえ多くの証拠が仮説が正しくないことを示しても、その仮説に基づいて、リジッド的な方法で行動する。
- 問題空間が大きくて、完全な探索ができない場合、あるいは時間が限られる場合、人は問題を解決できる最も可能だと思う仮説に基づいて行動する。
- また、ある特定の仮説を検証するために時間や精力を注ぐと、仮説を支持する証拠を見つける傾向が強くなり、反対の証拠についての処理が減少される。
- 複雑な問題を解決するためのもう1つの方略は、情報検索における encapsulation である。
- 人が問題を解決できない時、関連する情報を検索するのは役に立つかもしれない。

- しかし、新たに獲得した知識は、より多くの問題と不確実性をもたらす可能性があり、さらにより多くの情報検索行動を導く。
- 最後、情報を検索・獲得することは、一番の目標になってしまう。
- その結果、判断と行動が遅延されたり、行動が全然取られなかったりするようになってしまう。

1.2. Consequences of strategy use

- 非体系的な探索（試行錯誤、リジッド的な探索、encapsulation）を用いる際に、多くのエラーが起こることが先行研究により明らかになっている。
- エラーが発生した場合、人はタスクをコントロールできない感じがあり、自己効力感の低下に至る可能性がある。
- それに伴い、人は自分に対してよりネガティブな評価をする傾向がある。
- ネガティブな感情を処理するために認知資源が必要であるため、問題解決における使える認知資源が足りなくなってしまう、パフォーマンスの低下をもたらすと考えられる。
- そこで、本研究は、問題解決のプロセスでネガティブな自己評価（Negative self-evaluations）も分析する。
- 先行研究により、体系的な探索を使用する際に、感情の自己制御があることが明らかになっている。
- 本研究は、複雑な問題を解決するための異なる探索方略は、パフォーマンス（本研究ではコンピュータソフトウェアで解決できたタスクの数）に関連すると予測する。
- －（仮説1）具体的に、体系的な探索は、パフォーマンスとの間に、正の相関があり、非体系的な探索は、パフォーマンスとの間に、負の相関がある。

1.3. Error consequences

- 本研究では、問題解決の探索方略は、パフォーマンス以外に、エラーの数と種類にも関連していると考えられる。
- ヒューマンエラーに関する文献は、エラー自体とエラーによる結果を区別している。
- 次の行動に影響を与えるのはエラー自体ではなく、エラーの結果である。
- エラーの結果は、問題解決の方略とパフォーマンスとの間に重要な役割を果たしている可能性がある。
- 例えば、目標の達成にネガティブな影響を及ぼすエラーが発生した状況は、探索方略の変更を導く可能性がある。
- このようなエラーは、挫折を引き起こして、体系的な方略から非体系的な方略への切り替えをもたらす可能性がある。
- もう一方、エラーをして、意外に目標に近づくようになった場合もある。
- このようなエラーは、パフォーマンスに直接的にポジティブな影響を与える以外に、複雑な問題解決における重要である自尊心（self-esteem）にもポジティブな影響を与えて、パフォーマンスを間接的に向上させられる可能性がある。
- そこで、本研究は、エラーの結果に基づいて、エラーを3つの種類に分類する。
- －ネガティブなエラー（Negative error consequences）：目標の達成に直接負の影響を与える行為（例、次の有効な行動を妨害する行為または前できた結果を無駄にする行為）
- －ポジティブなエラー（Positive error consequences）：達成すべき目標に近づかないが、問題解決にポジティブな影響を及ぼす行為

- 影響なしのエラー (Non-effective actions) : 目標の達成に影響がない行為 (例、クリックできないボタンをクリックする行為)
- 仮説 : エラーとネガティブな自己評価
- ネガティブなエラーがたくさん起こる時、人はタスクをコントロールできないを感じて、自分が目標を達成する能力がないと感じる可能性がある。
- この場合、エラーはネガティブな感情を引き起こして、それをうまく処理できないと、次の行動の混乱に至り、さらに多くのエラーが発生すると考えられる。
- (仮説2) そこで、本研究は、エラーとネガティブな自己評価との間に強い相関があるという仮説を立てる。
- 仮説 : 探索方略とエラー
- 試行錯誤は、ランダムに行動を取るため、ネガティブなエラーと影響なしのエラーが多く起こると予測する。
- (仮説3) 試行錯誤は、ネガティブなエラーおよび影響なしのエラーとの間に、正の相関がある。
- リジッド的な探索は、ある特定の仮説にこだわり、問題における未知の情報空間を探索しないため、ネガティブなエラーとの相関がないと予測する。
- encapsulationは、情報検索に集中してしまい、積極的に行動を取れないため、ネガティブなエラーとの相関がないと予測する。
- (仮説4) リジッド的な探索とencapsulationは、ネガティブなエラーとの間に、相関がない。
- エラー、探索方略、ネガティブな自己評価の間に、具体的な相関関係についての予測は、表1 (“+” 正の相関、“-” 負の相関、“0” 相関なし)で示される。

Table 1. Expected relationships between strategies, performance and error consequences.

| | Performance | Positive error consequences | Negative error consequences | Non-effective actions |
|---------------------------|-------------|-----------------------------|-----------------------------|-----------------------|
| Systematic exploration | + | + | - | - |
| Trial-and-error | - | - | + | + |
| Rigid exploration | - | - | 0 | + |
| Encapsulation | - | - | 0 | 0 |
| Negative self-evaluations | - | - | + | + |

2. Method

2.1. Participants

- Amsterdam大学の心理学専攻の大学一年生19名(平均年齢=21.6, SD=5.5)、女性14名、男性5名

2.2. Task

- Excel(Macintosh4.0)で表のフォーマットを編集する課題 (テキストの移動、行の追加、色の変更、罫線の変更など8つのタスク)
- 課題は、エラー・トレーニングフレームワーク(error-training framework)の下で行われる。
- エラー・トレーニングは、トレーニング方法の一つであり、参加者に最小限の情報を提供して、エラーから学習させる。
- 全ての参加者もExcelの使用経験がないため、問題を解決するための探索は必要である。
- このようなトレーニングで、ほぼ全部の探索方略も観察されることができる。

2.3. Procedure

- 参加者が課題に取り組む前に、コンピュータ経験に関するアンケートを記入する。
- その後、5分間の短いExcel導入課題を行う。
- 導入課題で、参加者は実験環境と思考発話に慣れることができる。
- その後、参加者はExcelを操作して8つのタスクから構成される本番の課題に取り組む。
- タスクを行う方法についての教示はない。
- 時間制限は30分である。
- 参加者は、必ずしも時間内に全てのタスクを完了する必要はないが、できる限り多くのタスクを完了してくださいという指示を受ける。
- 実験中、実験者は参加者が見えないところで参加者を観察する。
- 参加者は5秒以上発話がない場合、実験者から「発話を続けてください」という指示を受ける。
- また、参加者の発話が録音され、コンピュータでの操作が録画される。

2.4. Measure

ー コンピュータ経験：

- OS、アプリケーション、ソフトウェアの使用経験に関する7つの質問に対して、5段階評定（1＝全くない、5＝よく利用）を行う(Reliability $\alpha = 0.75$)。
- 19名参加者のうち、14名は、Macの利用経験の得点が低く（得点＝1/2）、4名は得点が高かった（得点＝4/5）。
- 大部分の参加者は、windowsのOSの利用経験の得点が高かった（得点＝4/5）。

ー パフォーマンス：完了したタスクの数。

エラー、探索方略、ネガティブな自己評価についての分析方法（表2）

Table 2. Category descriptions, examples and interrater correlations.

| Category | Description | Examples | r |
|--------------------------------------|--|--|------|
| Systematic exploration | Exploring the system in a reflective manner: testing hypotheses, planning, and evaluation of action outcomes | - Considering a menu-option due to its name (e.g. border). - Analysing a menu-option accompanied with statements that suggest the processing of information available in that option. | 0.74 |
| Trial-and-error | Exploring the system without displaying signs of hypotheses testing, planning, or evaluation. | - Randomly pushing buttons or menu-options without sense of direction. | 0.87 |
| Rigid exploration | Repetitions of ineffective actions (more than twice), or holding on to faulty ideas longer than 'objectively' necessary. | - Keep on trying the option colour palette (more than twice) to colour a section without (any) effect. | 0.97 |
| Encapsulation in information seeking | Excessive search for information | - Staying in the help-menu for a long time and reading every bit of information carefully, irrespective of its use. | 0.93 |
| Negative self-evaluations | Statements that reflect perceived loss of control of the task | - 'I really cannot do this task'. - 'I have absolutely no idea what to do anymore'. | 0.77 |
| Positive error consequences | Erroneous actions that did not lead to the intended goal but that provide obvious useful information about the task. | - Accidentally changing the colours of a proportion of the table when trying to thicken the border. | 0.93 |
| Negative error consequences | Erroneous actions that led to consequences that had a direct negative effect on goal attainment. | - Deleting previous work. | 0.75 |
| Non-effective actions | Actions that did not have any consequences. | - Pressing an inactive button | 0.83 |

ー エラー：

- エンコーダーは、参加者がタスク中の操作をコーディングして、ある操作がエラーである場合に、その結果を評定して、ネガティブ、ポジティブまたは影響なしに分類する。

ー ネガティブな自己評価：

- 参加者の発話から、ネガティブな自己評価に関連する発話の回数を数える。
- 探索方略：
 - 参加者の発話に基づいて、タスク中の操作をコーディングして、探索方略を分類する。
 - エンコーダーは、ある操作が体系的な方略によるかまたは非体系的な方略による操作であるかを評定する。
 - もう一人のエンコーダーは、11名の参加者の操作をコーディングした。
 - 探索方略、ネガティブな自己評価、エラーによる結果についての分析結果に対して、二人のエンコーダーの間の相関を確認した。

各種類のエラーと探索方略の統計結果（表3）

- エラーと探索方略についての統計的な結果は、比率の形で表3で示される。
- 例えば、ある参加者は、タスク中に合わせて30個の操作をして、30個の操作のうち、10個はネガティブな結果をもたらすエラーであると、ネガティブなエラーの比率は0.33である。
- 探索方略も同じように、30個の操作のうち、10個の操作が体系的な探索による操作である場合に、体系的な探索の比率は0.33である。
- 探索方略を分析した結果、大部分の参加者は、タスク中複数の方略を使用した。
- 例えば、ある参加者はタスク中大部分の時間内で試行錯誤方略を用いて、時々体系的な探索を使用した。
- つまり、探索方略について、異なる参加者間で違うのは、使用した方略の種類ではなく、各種類の方略の使用頻度である。
- 参加者全員の平均を計算する。

Table 3. Participants' mean proportions, standard deviations, percentiles, and highest and lowest values for strategy-use and error consequences ($n = 19$).

| | M | SD | 25th percentile | 75th percentile | Lowest value | Highest value |
|-------------------------------|------|------|-----------------|-----------------|--------------|---------------|
| Strategies | | | | | | |
| Systematic exploration | 0.51 | 0.18 | 0.36 | 0.70 | 0.19 | 0.85 |
| Trial-and-error | 0.19 | 0.17 | 0.05 | 0.24 | 0.03 | 0.47 |
| Rigid exploration | 0.17 | 0.11 | 0.09 | 0.22 | 0.03 | 0.43 |
| Encapsulation | 0.03 | 0.05 | 0.00 | 0.07 | 0.00 | 0.25 |
| Error consequences | | | | | | |
| Positive error consequences | 0.05 | 0.08 | 0.01 | 0.11 | 0.00 | 0.21 |
| Negative error consequences | 0.18 | 0.05 | 0.09 | 0.25 | 0.05 | 0.40 |
| Non-effective actions | 0.44 | 0.11 | 0.38 | 0.51 | 0.31 | 0.67 |
| Actions not leading to errors | 0.33 | 0.14 | 0.21 | 0.41 | 0.00 | 0.48 |

Note: Proportions for strategies do not add up to 100% because not all actions that were coded could be indexed into one of the categories (rest category for strategies was 13%, there were no correlations with other study variables).

- 個人差が大きいですが、全ての参加者は体系的な探索、試行錯誤とリジッド的な探索を使用した。
- encapsulationは、使用が最も少ない方略であり、半分ぐらいの参加者がこの方略を使用しなかった。
- 各種類のエラーのうち、ポジティブな結果をもたらすエラーは最も少なかった。
- 20%ぐらいの参加者の操作で、ポジティブな結果をもたらすエラーはなかった。
- また、平均33%ぐらいの操作は、エラーでない操作であった。

3. Results (表4)

- 仮説を検証する前に、コンピュータ経験とパフォーマンスとの相関を分析した。
- その結果、コンピュータ経験とパフォーマンスとの間に、有意な相関関係が見られた。
- 探索方略がパフォーマンスに与える影響を分析するために、コンピュータ経験をコントロールした後、分析を行う。
- 探索方略 (4種類:No. 1-4)、ネガティブな自己評価 (No. 5)、エラー (3種類:No. 6-8)、パフォーマンス (No. 9)、操作数 (No. 10) の間の相関分析結果は表4で示される。
- 対角線の右上はコンピュータ経験をコントロールした上での相関結果、左下はコンピュータ経験をコントロールしていない結果である。

Table 4. Intercorrelation matrix strategies, error consequences, and task performance ($n = 19$).

| | M ^a | SD | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------------------------------|----------------|------|---------|---------|---------|---------|--------|--------|--------|---------|---------|--------|
| 1. Systematic | 21.84 | 7.5 | – | –0.62** | –0.60** | –0.57* | –0.49* | 0.73** | –0.51* | –0.71** | 0.59** | –0.45 |
| 2. Trial-and-error | 8.11 | 7.3 | –0.63** | – | –0.09 | 0.05 | 0.37 | –0.39 | 0.62* | 0.56* | –0.23 | 0.41 |
| 3. Rigid exploration | 7.26 | 5.0 | –0.65** | –0.02 | – | 0.61* | 0.09 | –0.56* | –0.13 | 0.29 | –0.52* | 0.02 |
| 4. Encapsulation | 1.42 | 2.2 | –0.60* | 0.11 | 0.60* | – | 0.02 | –0.47* | –0.19 | 0.14 | –0.60** | –0.15 |
| 5. Negative evaluation | 2.72 | 3.2 | –0.59* | 0.52* | 0.03 | –0.01 | – | –0.57* | 0.41 | 0.80** | –0.29 | 0.70** |
| 6. Positive consequence error | 2.17 | 1.9 | 0.77** | –0.37 | –0.61* | –0.51* | –0.56* | – | –0.42 | –0.81** | 0.65** | 0.60* |
| 7. Negative consequence error | 7.44 | 4.6 | –0.51* | 0.61* | 0.17 | –0.18 | 0.43** | –0.37 | – | 0.51* | –0.29 | 0.52* |
| 8. Non effective action | 18.83 | 6.5 | –0.73** | 0.56* | 0.32 | 0.16 | 0.85** | 0.79** | 0.48 | – | –0.40 | 0.85** |
| 9. Task performance ^b | 3.50 | 2.6 | 0.68** | –0.09 | –0.68** | –0.77** | –0.24 | 0.60* | –0.33 | –0.22 | – | –0.02 |
| 10. Number of actions | 42.47 | 12.9 | –0.45 | 0.44 | 0.05 | –0.14 | 0.93** | 0.75** | 0.95** | 0.60* | –0.13 | – |

Note: Correlations above the diagonal are zero-order correlations. Correlations below the diagonal are the partial correlations (controlled for general computer knowledge). ^aMeans denote the averages of the participants. ^bRange 0–8. * $p < 0.05$, ** $p < 0.01$.

3.1. Strategies and performance

- 探索方略とパフォーマンスとの相関を分析する。
- その結果、体系的な探索とパフォーマンスとの間に、正の相関が見られた。
- リジッド的な探索、encapsulationとパフォーマンスとの間に、負の相関関係が見られた。
- 試行錯誤は、パフォーマンスとの間に、統計的には有意ではないが、負の相関の傾向がある。

3.2. Error consequences and performance

- ポジティブな結果をもたらすエラーは、パフォーマンスとの間に正の相関がある。
- 影響なしのエラーとネガティブな結果をもたらすエラーは、パフォーマンスとの間に負の相関ある。
- 相関関係は仮説と一致しているが、有意な効果が見られなかった。

3.3. Strategies and error consequences

- 体系的な探索は、ネガティブなエラーおよび影響なしのエラーとの間に、負の相関があり、ポジティブなエラーとの間に、正の相関が見られた。
- それに対して、試行錯誤は、ネガティブなエラーおよび影響なしのエラーとの間に、正の相関があり、ポジティブなエラーとの間に、負の相関の傾向 (有意でない) が見られた。
- リジッド的な探索とencapsulationは、ポジティブなエラーとの間に、負の相関が見られた。
- また、ネガティブな自己評価は、ポジティブなエラーとの間に負の相関があり、ネガティブなエラーとの間に正の相関が見られた。
- さらに、操作数とネガティブな自己評価の間に、強い負の相関関係が見られた。

4. Discussion

- 分析結果により、仮説1の体系的な探索についての仮説は支持される。
- すなわち、体系的な探索はよいパフォーマンスとの間に正の相関がある。

- また、体系的な探索は、より多くのポジティブなエラー、より少ないネガティブなエラーとの間に正の相関がある。
- そのために、パフォーマンスと人間信頼性の面から見れば、体系的な探索は、複雑な問題を解決する際に最も期待される方略である。
- また、エラーについての分析は、エラー自体を厳密に分類するより、エラーの結果に基づいて分析するのは、有効な方法であることを示した。
- 仮説1の非体系的な探索についての予測と異なって、試行錯誤はパフォーマンスとの間に有意な負の相関が見られなかった。
- これは、試行錯誤を通して、人はタスクにとって必要なコンピュータ機能を探索できることを示した。
- しかし、仮説3の通り、試行錯誤は、ネガティブなエラーおよび影響なしのエラーとの間に正の相関がある。
- これは、試行錯誤を使用するのは、リスクとコストがあることを示した。
- また、リジッド的な探索とencapsulationは、パフォーマンスおよびエラーとの関係は同じような結果が示された。
- すなわち、リジッド的な探索もencapsulationも、パフォーマンスとの間に負の相関があり（仮説1の後半）、ポジティブなエラーとの間に、負の相関が見られた。
- しかし、仮説4と一致して、リジッド的な探索もencapsulationも、ネガティブなエラーおよび影響なしのエラーとの間に、相関が見られなかった。
- また、仮説2は支持されて、パフォーマンス、ネガティブなエラー、影響なしのエラーは、ネガティブな自己評価との間に、正の相関がある。
- 本研究は、相関関係を分析するため、ネガティブな自己評価についての因果関係を出すことができない。
- しかし、観察により、ネガティブな自己評価は、よくネガティブなエラーが発生した後、または長い間に影響なしのエラーが続いた後生じる。
- さらに、ネガティブな自己評価と、操作の数との間に強い相関が見られた。
- これは、先行研究で言及された緊急状態における反応（emergency reaction）と類似していると考えられる。
- つまり、人はあるタスクをコントロールできないと感じて、自己効力感の低下が起こる際に、積極的に行動を取る。
- このような行動（非体系的な方略）はある程度役に立たない可能性がある。
- しかし、積極的に行動を取るのは、自分がまだコントロールをしている印象を作って、複雑な問題を解決する際に必要な自尊心を守るためである。

4. Future study

- 今後の研究では、複雑な問題を解決するプロセスで、方略の変更はどのように起こるか、またネガティブなエラーは方略の変更を引き起こすかをさらに検討する。
- そして、自己制御と探索方略との関連を具体的に検討する。