

A Model of the Perceptual and Conceptual Processes in Graph Comprehension

Carpenter, P.A. and Shah, P.

Journal of Experimental Psychology: Applied, 4, 75-100.

- はじめに
 - 本研究では、線グラフを理解のモデルを提案する
 - ◇ 提案モデル…ラベルや尺度などの解釈を伴う概念的なプロセスと、ラインの傾斜やパターンのエンコードなどの知覚的なプロセスの緊密なインタラクションを強調するモデル

- Task Analysis
 - 本研究はBertin(1983)によって初めて詳述され、その後多くの研究者によって改良がくわえられてきた、グラフ解釈の課題分析に基づいている (e.g., Cleveland & McGill, 1985; Guthrie et al.,1993; Lohse, 1993; Pinker, 1990; Simkin & Hastie, 1987)
 - ◇ この分析はグラフ理解のプロセスとして3つの主要なプロセスを示している
 - 視覚的なパターンのエンコード
 - パターンと、概念的な関連もしくは量的な関係との間の翻訳
 - 指示されたもの (ライン) の関係の同定
⇒ パターン認識モデル
 - もう一つのモデル：統合モデル (integrative model)
 - ◇ グラフは視覚的なチャンクに分解され、視覚的なパターンと量的・質的な解釈が統合される
 - 具体的には線グラフのライン
 - ◇ パターンと解釈の統合はチャンクごとに繰り返される

- Graphic Complexity
 - 本研究では3つの連続型変数の線グラフを使用する
 - ◇ 3変数のデータセットは交互作用を含むような複雑性の操作が可能なので、本研究の課題として適している
 - y が従属変数、独立変数の1つが x 、もう一方が z ラベル (凡例)
 - ◇ 著者らの先行研究では、 x - y 関係がひとまとまりとしてエンコードされ、解釈されるという仮説を支持する結果となった (Shah & Carpenter, 1995)
 - 解釈の際、 x - y 関係の数量的な情報はよく叙述されるが、 z - y 関係のラベル間の量的 (metric) な関係にはほとんど触れられない
 - Figure1A と Figure1Bは、もとは同じデータセットから作られたグラフ
⇒互いに相容れない表現だと判断される (元が同じデータだとは気付かない)

- ◇ 同じデータをもとにしたグラフでも、ライン同士の向きが変わることがある。このグラフ間で処理時間などに違いがあれば、関数の組み合わせの複雑さがグラフ理解のプロセスに影響していると考えられる
 - Fig.1Aは2本のラインの傾きは似ているが、Bは逆の角度のラインであり、より複雑である
 - ◇ ラインの密度の高いグラフと密度の低いグラフを使用して、グラフの複雑性を操作する
 - Fig.1Cに加えられたラインはすべて似たような角度だが、Dに加えられたラインは角度がバラバラで、より複雑なグラフとなっている
- **Cross-over Interaction データセット(Figure1A/B)**
- ◇ グラフの視点：Cross-Over グラフ / Non-Cross-Over グラフ
 - ◇ 交差グラフはどちらのラインも増加している。xの主効果が際立っている
 - ◇ 交差しないグラフは片方のラインは増加し、もう一方のラインは減少している
 - ⇒ 交差しないグラフの読み取りは複雑になる
 - ◇ 交差グラフにおいてはラインの密度が高くなっても、注視の時間と回数にほとんど影響はない
 - ← 密度の低いグラフと基本的な傾向は変わらないから
 - ◇ 交差しないグラフにおいては、ラインの密度が高くなると逆の角度のラインを埋める角度を持つことになり、付加的なエンコーディングと解釈の過程を必要とする
- **Acceleration データセット(Figure2A/B)**
- ◇ グラフの視点：Curve グラフ / Intercurve Spacing グラフ
 - ◇ Fig2A の Curve グラフでは加速度的増加（減少）は見えやすいが、Fig2B の Intercurve Spacing グラフでは見えにくい
 - ◇ 密度の操作においては、どちらの視点のグラフにおいても、加わるラインはもとのラインと平行になるので、密度の効果はほとんどないと考えられる
- **Common Value Interaction データセット(Figure3A/B)**
- ◇ グラフの視点：Line グラフ / Point グラフ
 - ◇ ひとつの独立変数の1水準において、他方の要因の単純主効果がないデータセット
 - ◇ Fig3A の Line グラフでは一つの独立変数の1水準において他方の要因の単純主効果がないということが見えやすいが、Fig3B の point グラフでは見えにくい
 - ◇ 密度の効果は Cross-over Interaction データセットと同様にグラフの視点の違いによって効果の大きさが異なる
 - Point グラフ：効果小 / Line グラフ：効果大

- 実験 1

- 目的

- ◇ 線グラフの注視時間を検討して 2 つのモデルから予想される結果とどちらがあっているかを検討する
- ◇ パターン認識モデル (pattern-recognition model)
 - 軸と z ラベル (凡例) に対する注視時間と回数は少ない
 - 線のメインパターンに対する注視時間が長い
- ◇ 統合モデル (integrative model)
 - パターンのエンコードとラベルなどの解釈との統合の循環的なモデル
 - x 軸, y 軸, z ラベルの解釈の時間が長い
- ◇ 先行研究と同様の結果 (x-y 関係の叙述 > z-y 関係の叙述) が出るかも確かめた.

- 方法

- ◇ 課題

- 12 ペアのグラフ
- 初めのグラフ読み取りについて発話
- 眼球運動測定
- 比較課題
 - 初めのグラフと対のグラフは矛盾しているか, していないか (incompatible / compatible)

- ◇ 実験参加者

- カーネギーメロン大学の学生 40 名
 - ほかに 6 名は眼球運動が正確に記録できなかったため分析から除外

- ◇ 実験計画

- 実験 1
- 3 要因混合計画
 - 第一被験者内要因 : データ領域 3 種
 - 第二被験者内要因 : グラフの視点 (x 軸と凡例の配置)
 - 被験者間要因 : グラフのデータセットの密度 (高い / 低い)

- ◇ 実験刺激

- 密度高のグラフ 12 ペア, 密度低のグラフ 12 ペア (Figure1~3 参照)
 - 3 変数のグラフ
 - ◇ 関連がありそうだが, 具体的な関係がよく知られていないもの
ex.) 語彙スコアと年齢・TV 視聴時間
 - 比較課題用刺激
 - ◇ 3 変数のグラフの x 軸と凡例を入れ替えたもの (正答用)

◇ 上記と少し違うもの（誤答用）

- 比較課題の正答用と誤答用それぞれと組み合わせたペア

◇ 手続き

- トライアル
 1. 初期グラフが提示される
 2. 初期グラフについての記述（読み取り）をする
終了したらボタンを押す
 3. 比較グラフが提示される
「2つのグラフは同じデータセットの異なる視点のグラフである可能性があります」
- グラフセット
 - 練習用 1 ペア
 - 本番用 12 ペア
 - フィラー 5 ペア
- グラフの z 変数は名称的や順序的ではなく、連続的であるという教示を行った
 - z 変数の **metric** な特徴のエンコードの失敗が、順序や名称的な特徴であるという間違っただ推論を反映しないように

◇ Eye Fixation

- Iscan RK-426 使用
- 16.7ms 毎に xy 座標を出力
- 画面は 5つの領域に分割された(Figure4 参照)
- 250ms 以上の注視を分析対象とした
 - ノイズやまばたき、サッカード(saccade)の影響を除外するため
 - この基準で注視時間の 10.7%が分析から除外
- 30 試行分のデータが分析から除外された
 - 15%以上のデータが失われた試行 (16 試行)
 - 75 秒以上かかった試行 (8 試行)
 - 実験参加者が飛ばしてしまった試行 (6 試行)
 - 除外されたデータは平均値で補われた

◇ data analysis

- Table1 は総注視時間と 3つのデータセットごとの注視時間
 - 著者らの興味は視点と密度が異なる 4種のグラフタイプの比較にあるので
- それぞれのデータセットの中で ANOVA を行った
 - 独立変数は密度と視点／従属変数は注視時間と注視回数

- Table1 には F 値が 2.5 以上だったものを示している
 - ◇ 数値と*印は注視時間, §は注視回数において有意であることを示している
- ◇ 口頭記述の分類
 - 初期グラフについての叙述は, それぞれのグラフの z ラベル (凡例) についてどのような叙述を行ったかによって分類された
 - (a) nominal descriptions : 名称的な叙述
 - (b) ordinal descriptions : 順序的な情報を含む叙述
 - (c) metric descriptions : 量的関係を含む叙述
- 結果と考察
 - ◇ 課題遂行時間と注視時間
 - 全体
 - 平均注視回数 : 33 回 / 1 試行
 - 平均注視時間 : 1.1 秒 / 注視 1 回
 - ◇ 最小 : 29 回 (33 秒) (intercurve spacing graph)
 - ◇ 最大 : 44.5 回 (40.6 秒) (non-cross-over dense graph)
 - Figure5 は典型的な視線推移のパターン (Figure1C)
 - グラフパターンと各ラベル, タイトルの間の推移が多い
 - ← 水平方向・垂直方向の推移が多いことから
 - 流れとしては, まずタイトルを読み, グラフをさっと見て変数間の関係を述べる

典型的な回答例「これは『テレビ視聴時間』の違いによる『語彙スコア』と『年齢』の関係で..., テレビをたくさん見る人においては年齢とともに語彙スコアが劇的に増加するが, テレビを少ししか見ない人にとっては, 年齢による語彙スコアの増加はそんなに劇的ではない」
 - この視線推移のパターンは, グラフ上の異なる領域の情報を組み合わせるには一度リフレッシュしてから統合する必要があることを示唆している
 - Figure6 は 5 つの領域ごとの平均注視時間
 - 統合モデルによって予想されたように, ラベルと変数の値にかなりの時間注視している
 - ◇ グラフパターンが平均 9 秒であるのに対し, ラベルや変数は計 24 秒も
 - ◇ タイトル注視時間の 1-2 秒は初めの読み上げ時間と考えられる
 - このことは, 「人はおもにグラフパターンに焦点をあてる」という仮説に反する

- 遂行時間の結果はグラフの密度と視点の効果を示していた (Table 1 参照)
 - 統合モデルに基づく予想に合致して, x-y 関数 (ライン) の数が増えると処理時間も増加していた (ライン同士が平行でないグラフにおいて)
 - 同じデータセットから作られたグラフであっても, 視点が違うと処理時間に有意差があった
 - ⇒ 同じデータからであっても視点が異なると, 異なる心的表象を作り出すことを示唆している
 - 統合モデルに基づく予想では, 密度が高くなると処理時間が増加する. 特に明らかに異なる関数の (=ライン同士の角度が逆の) ラインのグラフにおいて
 - Figure7 にグラフごとの結果を図示した
 - ◇ cross-over interaction データセットは密度×視点の交互作用が有意 : $F(1,32)=5.82, p<.05$
 - 特に, 平行に近いラインが加わる Cross-Over 視点よりも平行でないラインが加わる Non-Cross-Over 視点のほうが, 密度による処理時間の増加が顕著であった
 - ◇ common-value interaction データセットでも同様の傾向だった (有意傾向) : $F(1,32)=5.82, p=.051$
 - ◇ acceleration データセットにおいては, 加わるラインがどちらの視点でも平行なので, 統合モデルからは, 交互作用が有意にはならない → 有意ではなかった : $F(1,32)=0.62, n.s.$
 - ⇒ 明らかに異なる関数的な関係は解釈の時間に影響を及ぼす
- 各グラフの領域ごとの注視時間における視点と密度の効果 (Table1 参照) ーどの領域の差が全体の差につながったのか
 - cross-over interaction データセット
 - ◇ グラフパターンにおいてもラベルにおいてもほぼ全てで交互作用が見られた
 - common-value interaction データセット
 - ◇ グラフパターンと z-ラベルにおいて交互作用が見られた
 - ⇒ おそらくこのデータセットのグラフではこの 2 つの領域において追加されたラインの処理が増えるから
 - acceleration データセット
 - ◇ 視点の主効果のみ
 - 全体の時間, グラフパターン, x-軸においては curve

perspective の方が有意に長い

- z-ラベルのみ intercurve spacing perspective の方が有意に長い

- 遂行時間と注視時間の結果から、「グラフパターンの認識→（ラベルの？）解釈→統合→グラフパターンの認識」の循環的な処理が確認された
- チャンクに分解して解釈するということから、以下に述べる「z-y関係の metric な特徴の理解が不足する」ということを補助している

☆ Descriptions task performance

- x-y関係を叙述することは比較的簡単で、z-y関係の metric な推測を行うことは比較的難しい
 - 個々の z の水準における x-y の metric な関係についてはよく叙述されるが、z 変数については、nominal もしくは ordinal な情報しか含まれない
 - Figure8 は各グラフにおける z-y 関係についての叙述のカテゴリごとのヒストグラム

☆ 明らかに nominal の叙述が多い

☆ ordinal と metric の叙述の頻度には視点による違いがある

- グラフパターンについての予想通り，
 - ✓ z-y 関係の ordinal もしくは metric な叙述は平行な x-y ラインを持つ curve グラフと intercurve spacing グラフに多かった。
 - ✓ 逆に、ライン同士の角度がまったく異なる（逆方向の）グラフである non-cross-over グラフにおいて、nominal 以外の叙述が最も少なかった。
- ラインの密度についての予想通り，
 - ✓ 密度の違いによって、叙述の割合に有意差があった
: X^2 乗(1,408)=22.9, $p < .01$
 - ✓ 密度が高い＝ラインがたくさんあると z における各水準の間の量的な関係に注目しやすくなるから
 - ✓ intercurve spacing グラフにおいてのみ、密度が低い方が metric な叙述が多い
← ラインが多くなるとライン間の不均等な空間が見えづらくなるから

☆ Comparison task performance

- 実験参加者が初期グラフの z-y 関係について正しい表象を持っていたかを測定するために実施
- 全般的に実験参加者はあまり正しい判断ができなかった

: 平均正答率 52.2%

⇒ 多くの実験参加者は z-y 関係について正しい表象をもっていない

- metric な叙述をしていた参加者は正答率が高かった
: metric: 74.2% ⇔ nominal:50.8%, ordinal:50.6%
- 密度の効果はなかった: 高密度:55.9%; 低密度:28.5
: $F(1,32)=1.58, n.s.$

➤ まとめ

- ◇ 異なる視点のグラフにおける注視時間と回数の結果から統合モデルが支持された
 - パターンの認識と翻訳と同定のプロセスは、グラフパターンと特定のラベルや軸と関連付ける視線推移によって引き起こされる
 - グラフの解釈はパターン認識と統合の漸進的なプロセスによって起こる
 - しかしその統合は、z-y 関係よりも x-y 関係において多く観察される

● 実験 2

➤ 目的

- ◇ 実験 2 の目的は、注視パターンを代数的なモデルにフィットさせることによって、解釈のプロセスをより詳細に特徴づけることである。
 - 統合モデルが正しいなら、平行でないラインが増えるに従って、それぞれの領域における注視回数は単調に増加するはずである
 - 増加したライン 1 本 1 本について統合プロセスが起こるから
- ◇ 視線と口述の関係を組織的に検討するために、グラフについて実験者からの質問に回答してもらうようにした
 - Figure1-A に対する質問例
 - x-y 関係を訊く質問「年齢が増加するに従って、語彙スコアはどのようになっていますか」
 - z-y 関係を訊く質問「テレビの視聴時間が増加するに従って、語彙スコアはどのようになっていますか」
- ◇ 質問に答えてもらう形にすることで、x-y 関係を叙述中の視線と z-y 関係を叙述中の視線のパターンの違いを検討することができる
 - なぜ z-y 関係を読み取るのは難しいのか

➤ 方法

◇ 実験参加者

- カーネギーメロン大学の学生 29 名
 - ほかに 4 名は眼球運動が正確に記録できなかったため分析から除外

◇ 実験計画

- 2 要因混合計画

- 第一被験者内要因：データ領域 3 種
- 第二被験者内要因：視覚的な見方
- 半数の参加者は 1 つのグラフに対して x-y 関係の質問と z-y 関係の質問の両方を受ける
- 残りの参加者は 1 つのグラフに対しては、どちらか一方の質問のみ
- ◇ 実験刺激
 - 比較課題は行わないので入れ替えグラフはなし
 - 密度高のグラフ 6 個，密度低のグラフ 6 個
- ◇ 手続き
 - 基本的に実験 1 と同じ
 - 違う点
 - はじめに実験者がグラフのタイトルを読み上げて，質問をする
 - ◇ そのあと実験参加者は注視点をみてボタンを押してグラフを提示させる
 - 250ms 以上の注視を分析対象とした
 - 15%以上のデータが失われた 16 試行分のデータが分析から除外された
- ◇ The coding scheme
 - それぞれのグラフについて，固有の x-y 関係の数をあらわした値を付与した
 - 1 本のラインは “1”
 - それに対して平行なラインが加わっても値は変わらない
 - 同傾向だが角度が異なるラインが加わると 0.5 増加で “1.5”
 - point グラフ，cross-over グラフ ※密度高・低にかかわらず
 - 逆向きのラインが加わると “1.75”
 - line グラフ
 - 直角に交わったラインが加わると “2”
 - cross-over グラフ（密度低）
 - cross-over の密度高では “3”
 - 幅は 1 ~ 3
- 結果と考察
 - ◇ 質問回答中の総時間と注視回数の分布の結果は統合モデルを支持した
 - 全グラフの平均時間：23 秒
 - 最短平均：intercurve spacing グラフ（17.8s）
 - 最長平均：non-cross-over グラフ（31s）
 - Figure9 はそれぞれの領域に費やした時間
 - ◇ 実験 1 と同様に，グラフパターンの時間よりも，ラベルや各軸に費やした時間の方が長かった

☆ グラフパターンが平均 6.3 秒であるのに対し、ラベルや変数は計 14.4 秒も

- 全グラフの平均注視回数：23.5
 - 最小：intercurve spacing グラフ（18.4 回）
 - 最大：non-cross-over グラフ（32.6 回）

☆ Comprehension of x-y relations.

- x-y 関係の解釈を組織的に分析するために、シンプルな代数モデルを適用して、グラフタイプごとの領域別注視回数の予想を立てた
- グラフごとの注視の分布は大きく分けて 2 つのプロセスによって決定した

➤ グラフのスキヤニングプロセス

(a) グラフの概観と質問が指示しているものの位置を取得する

(5 つの領域を一度に見る)

(b) 一つの関数（ライン）を同定する

(c) その関数の x-y 関係の性質を叙述する

1. x 軸における変化（増加か減少か）を同定する

- 方向，スケール，構成単位と指示対象の情報も含む
(x 軸を見る→パターンを見る→x 軸を見る)

2. y 軸における変化（増加か減少か）を同定する

- 方向，スケール，構成単位と指示対象の情報も含む
(y 軸を見る→パターンを見る→y 軸を見る)

3. z-ラベルを同定する

- 方向，スケール，構成単位と指示対象の情報も含む
(z ラベルを見る→パターンを見る→z ラベルを見る)

☆ (b)と(c)のプロセスは異なる関数（ライン）に繰り返して適用される

☆ (c)の 1~3 の順番はその時々で変わる

ex)ある特定のラインを解釈しようとするときは、まず最初に z ラベルを見て同定する

➤ 人はグラフパターンの照査を挟んでそれぞれのラベルを 2 回見る

☆ 複数回の照査はいくつかの下位コンポーネントを反映している

- 情報の確定，新しい情報のチェック，関連する情報に対する記憶のリフレッシュ

☆ これらのことは一回の注視でエンコードしたり保持したりできる情報に厳しい限界があるという仮定と整合する

- 全グラフにおける、実際の平均注視回数と予想回数を Table2 に示した

➤ おおよそ予想できている

☆ 個々のグラフごとの結果をプロットした Figure10 でも同様である

- ◇ パターン領域における実際の注視回数が予想を上回ったのは、指示された関数（ライン）の同定のためと思われる
 - ◇ タイトル領域における実際の注視回数が予想を上回ったのは、いくつかの変数の同定のためにタイトルを確認したためと思われる
 - 異なる関数の数が理解のプロセスに影響を与えているということは、総注視回数の高い相関によって支持された
 - ◇ 個別の領域においても .77～ .90 という高い相関
 - 視覚的な複雑性の評定から作ったモデルとの比較
 - 学生 10 名が個々のグラフの複雑性を評定
 - 評定結果をもとにした代数モデルで注視回数を予想
 - ◇ 領域全体の注視回数においては、部分的には（グラフによっては？）高い相関を示すものもあったが、全グラフでは $r(10) = .43, n.s.$
 - ◇ 個別の領域においては負の相関になるものもあった $r(10) = -.67$
 - グラフの複雑性評定はラインの本数が多いとより複雑であると判断される傾向にあり、異なる関数の数を反映していないから
- ◇ Graph-type analysis.
 - これまでの結果から、同じデータセットから作られたグラフであっても異なる内的表象を引き起こすということがわかった
 - Table3 は x-y 関係について訊かれた時の注視時間の平均の一覧、Figure11 はそのグラフである
 - 予想通り、平行でないラインが加わった non-cross-over グラフにおいて、平行に近いラインが加わる cross-over グラフよりも密度の効果が大きく出ている ($F(1,27) = 8.74, p < .05$)
 - line グラフと point グラフにおいては交互作用は有意ではなかった
 - グラフの視点、密度に質問の種類を加えた 3 要因の分散分析をしたところ、2 次の交互作用が有意であった ($F(1,27) = 3.74, p < .05$)
 - 実験 1 と同様に、curve と intercurve グラフにおいては交互作用は有意ではなかった
 - これらの結果は統合モデルを支持する
- ◇ Density manipulation
 - 加わるラインが固有の関数を持っているときのみ密度の効果が出る
 - ← 密度の操作はパターン領域と z ラベル領域においてみられる
 - ← ラインが増えると個々のラインと名称を関連付ける作業が必要だから
 - 密度の高い低い比較
 - パターン領域において
 - ◇ 7.4s vs. 5.1s ($F(1,27) = 4.66, p < .05$), 回数 : 8.1 vs. 6.3

