

Cognitive Load Measurement as a Means to Advanced Cognitive Load Theory

Fred Paas, Juhani E. Touvienen, Huib Tabbers, and Pascal W. M. Van Gerven

● INTRODUCTION

- ✓ Cognitive load theory (CLT) の関心・目的
 - ◇ 人間の限られた認知資源を効果的に利用するインストラクション方法の発展
 - 理論の基礎となる認知アーキテクチャ
 - 限られたワーキングメモリ
 - ⇒理論の中心
 - 無限の長期記憶

- ✓ CLT の最も重要な目的：スキーマの構築と自動化
 - ◇ スキーマの構築・自動化のためには適切な認知資源の配置が必要
 - 負荷がかかりすぎても、少なすぎてもパフォーマンスが悪化
 - ⇒教授者があらかじめ適切であろう負荷を想定し、効果的な学習環境を構築することが重要

- ✓ これまでの研究の結果
 - ◇ 様々な年齢、様々な学習領域でその効果を確認
 - CLT-base の課題の特徴
 - トレーニング時間の短縮
 - 少ない mental effort での学習 (mental effort については後述)
 - 通常の学習と同程度 (もしくは勝る) の学習効果

- ✓ 本稿における議論
 - ◇ 認知負荷の測定が CLT の成功に貢献するということと、さらなる発展のために不可欠なものであることを論じる

● THE CONSTRUCT OF COGNITIVE LOAD

- ✓ Cognitive load の定義 (Paas & van Merriënboer, 1994a)
 - ◇ 特定の課題において、学習者の認知システムに課される負荷。多次元の要素から構成
 - 課題の特徴と学習者の特徴の相互作用により量が決定される
 - 課題の特徴の例
 - ◇ フォーマット、複雑さ、マルチメディアの使用法、タイムプレッシャー etc.
 - 学習者の特徴の例
 - ◇ 熟達度、年齢、空間能力

- 相互作用の例
 - ◇ フォーマットと熟達度の相互作用
 - 初心者の頃は有効であったフォーマットが、学習が進むと有効度が減少、または逆効果になってしまうことがある(*expertise reversal effect*)

- どのようなものが測定可能か?
 - *Mental load*
 - ◇ 課題と対象の特徴の相互作用により決定されるもの
 - 課題, 対象の特徴に対する知識により決定(Paas & vanMerriënboer, 1994a)
 - あらかじめ予測可能
 - *Mental effort*
 - ◇ 実際に割り当てられる認知容量
 - 実際の認知負荷と考えられる
 - *Performance*
 - ◇ 認知負荷の一側面
 - 正解数, エラー, 課題にかけた時間などで測定

測定可能

✓ 認知負荷の特性と定義のフレームワーク

◇ CLT における 3 つの cognitive load

- *intrinsic load*
 - 学習者の熟達度と学習教材の相互作用により決定
 - ◇ 学習者が現在持っている知識・熟達度が影響
 - ⇒教材作成者が直接操作することはできない

- *extraneous or ineffective load*
 - 良くないデザインが原因となる, 必要のない認知負荷
 - ◇ 図と, それを説明するテキストが離れて配置されていると悪影響
 - ⇒教材の善し悪しが直接的に影響する: 教材作成者が管理可能なもの

- *germane or effective load*
 - スキーマ構築, 自動化の構築に関連するもの
 - (他の研究では **Self-explanation** が挙げられている)
 - ⇒教材作成者が直接操作することはできない

※CLT の考えでは, 上記 3 つは加算的

⇒overload にならないようにこの 3 つの負荷を管理しなければならない

◇ Xie らによる概念的フレームワーク：図 1

- *instantaneous load*
 - 瞬間々々における認知負荷
- *peak load*
 - 課題中における負荷の最大値
- *accumulated load*
 - 課題中の負荷の総和
- *average load*
 - 負荷の平均強度
- *overall load*
 - 課題全体を通じた負荷

⇒これらは概念的理解に役に立つ。また、これらの指標は生理学的手法にて視覚化される。

● MEASUREMENT OF COGNITIVE LOAD

✓ 認知負荷の測定の困難さ…多くの研究で示される

◇ 負荷測定の代表的な方法

- 解析的手法
 - 数学モデルや課題解析
 - ◇ プロダクションシステムによるシミュレーション
- 実験的手法
 - 評定用紙，課題パフォーマンス，二重課題法，生理学的指標

✓ これまでの研究において用いられてきた方法一覧：表 2

◇ 圧倒的に実験的手法の方が多し：なぜかは後述

✓ 測定法の概観

◇ 評定用紙

- 人は *mental effort* をきちんと評定できるのだという報告のもとに使用 (Gopher & Braune, 1984)
- 具体的な評価法 (最も用いられている方法)
 - *Mental effort*
 - ◇ 9 (7) 段階で *very high effort, high effort, ..., low effort, very low effort* を評定
 - ◇ 相対的に小さな *cognitive load* の変化にも敏感であり，信頼性も確認されている (Gimino, 2002, Paas et al., 1994)

◇ 生理学的手法

- 認知機能の変化は，生理学的な変化に反映されるという仮定のもとに使用
- 具体的な評価法
 - 心拍，脳活動，眼球運動など
- 利点
 - 負荷の傾向やパターンを視覚化可能 (表 1 のような表現ができる)

- 実例
 - 心拍を使った実験(Paas and vanMerriënboer,1994b)
 - ◇ 煩わしく，無効果，認知負荷の変化に対する反応が鈍い
 - TEPR(Beatty and Lucero-Wagoner, 2000)
 - ◇ 瞳孔の反応により評価：効果的である
- 問題点
 - 機器の問題（高い，拘束性がある），信頼性の問題
 - ⇒だから，あまり用いられていない

◇ パフォーマンス

- 具体的評価法
 - 2重課題法
 - ◇ 主課題と副課題を同時に遂行
 - 副課題の成績が認知負荷の指標となる
 - 測定項目の例：反応時間，正確さ，エラー数
 - ⇒副課題の成績が悪い＝主課題で一杯一杯＝主課題は過負荷
 - ◇ 具体例
 - ランダム文字生成（副課題）を行いながらゴルフパットの練習（主課題）
 - 問題点
 - ◇ 主課題が複雑な時に，特に主課題に与える影響が大きい
 - ⇒だから，あまり用いられていない

● MENTAL EFFICIENCY OF INSTRUCTIONAL CONDITIONS

- ✓ 認知負荷の測定と成績の評価は，異なる学習条件の効果評定に重要だと考えられるが，それだけでは駄目
 - ◇ なぜ？
 - 人は課題が複雑になった時，より mental effort をかけることで成績を維持可能
 - それぞれを単体で考えるのではなく，統合的に考える必要がある
- ✓ Paa らのアプローチ
 - ◇ 基本的な考え
 - Low-effort で high performance : high-instructional efficiency
 - High-effort で low performance : low-instructional efficiency
 - ◇ 算出方法
 - performance, mental effort をそれぞれ z-score(平均 0, 標準偏差 1)に標準化
 - instructional efficiency(E)を以下の式で算出 (図 2 参照)

$$E = \frac{Z_{Performance} - Z_{MentalEffort}}{\sqrt{2}}$$

- 図2の概要(3条件(A, B, C)のEがそれぞれ図中の位置にプロットされたとする)
 - A: 相対的に低い mental effort で好成绩⇒最も効率が良い
 - ◇ 成績が高いが, mental effort は低いと評定
 - B: 中間的な効率
 - ◇ 成績, mental effort とともに中間
 - C: 高い mental effort で低成績⇒最も効率が悪い
 - ◇ 成績は低い, mental effort が高いと評定

● OVERVIEW OF STUDIES

- ✓ 表1: 認知負荷の測定を行った研究の概要

- ◇ 興味深い点

- 生理学指標, 二重課題の手法を発展させようという試みはなされているが, 未だ9(7)点法の測定法が根強い
 - ⇒簡単に使用できるし, 本課題に影響を及ぼさない, 高価でない, 感受性が高い, 信頼性がある. しかし, 強く主張しておきたいことは, 内部的な一貫性についてはさらなる研究が必要であるということである.

● DISCUSSION

- ✓ 認知負荷の測定は CLT の発展にどう貢献するのだろうか?

- ◇ 認知負荷が教育へ与える影響の実証的基礎を構築する

- 具体例

- Worked-out example の使用

- ◇ テストパフォーマンスが良いだけでなく, extraneous load が少ない

- Split-attention effect など

- ◇ フォーマットの違いにより, ワーキングメモリの使用量が異なる

⇒中には失敗した例もあるが, 大部分の研究はフォーマットの違いによる memory load の違いが学習効果に影響することを示す

- ◇ mental efficiency

- 教材の効果だけでなく, 効率も測定できる

- 多くの研究で効果は同じでも効率が異なることが示されてきている

⇒mental efficiency の検討は, フォーマットの違いが学習プロセスの違いに与えるという我々の知識を豊かにする

- ✓ 評定尺度法(9(7)点法による評定)と mental efficiency についての注意喚起

- ◇ とても有用で使いやすい (ので多くの研究で用いられている)

- 起源は Paa ら(1992,1994) によるものであり, それらの研究では mental effort を評定

- 追従する研究では, 異なるスケールや, 課題の困難さなど異なるものを評定している

⇒これらについては, 信頼性が得られている訳ではないのでさらなる検討が必要

- ✓ 今後に向けて
 - ◇ 要検討課題
 - 課題にかけた時間
 - mental effort を 5 と評定された時に、課題にかけた時間が 5 分の時と 10 分の時ではどのような違いがあるのだろうか？
⇒時間を考慮に加えた検討が必要
 - load の違いを明確に
 - 何が intrinsic load, extraneous load, germane load なのかを明確にする必要性
⇒解析的手法と実験的手法を組みあわせて検討していく必要
 - ◇ 新たな生理学的指標の模索
 - PET, MRI
⇒異なるタイプの負荷の神経基質的基盤を築いてくれるだろう（まだまだ高価だけど...）
 - ◇ 応用の可能性
 - mental effort と performance を使用した、個別化知的インタラクティブ学習支援システム
 - 個人の違いに対応し、適切な学習課題を課すシステムが出来る！

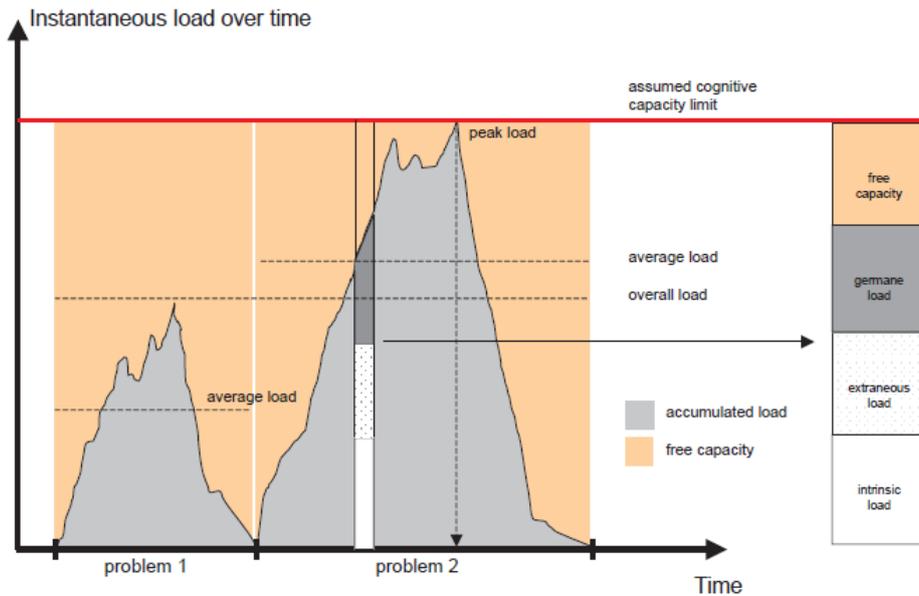


FIGURE 1 Attributes of cognitive load and a framework of cognitive load definitions.

TABLE 1 Studies That Measured Cognitive Load and Calculated Mental Efficiency and the Measurement Technique They Used

Studies	Cognitive Load Measurement Technique	Mental Efficiency
Sweller (1988)	PS, ST	
Paas (1992)	RS9	
Paas & van Merriënboer (1993)	RS9	ME
Paas & van Merriënboer (1994b)	RS9, HRV	ME
Cerpa, Chandler, & Sweller (1996)	RS9	ME
Chandler & Sweller (1996)	ST	
Marcus, Cooper, & Sweller (1996)	RS7, ST	ME
Tindall-Ford, Chandler, & Sweller (1997)	RS7	ME
Yeung, Jin, & Sweller (1997)	RS9	ME
de Croock, van Merriënboer, & Paas (1998)	RS9	
Kalyuga, Chandler, & Sweller (1998)	RS7	ME
Kalyuga, Chandler, & Sweller (1999)	RS7	ME
Tuovinen & Sweller (1999)	RS9	ME
Yeung (1999)	RS9	ME
Kalyuga, Chandler, & Sweller (2000)	RS7	ME
Kalyuga, Chandler, & Sweller (2001)	RS7	ME
Kalyuga, Chandler, Tuovinen, & Sweller (2001)	RS9	ME
Mayer & Chandler (2001)	RS7	
Pollock, Chandler, & Sweller (2002)	RS7	ME
Stark, Mandl, Gruber, & Renkl (2002)	RS9	
Tabbers, Martens, & van Merriënboer (2002)	RS9	
Tabbers, Martens, & van Merriënboer (in press)	RS9	
Van Gerven, Paas, van Merriënboer, Hendriks, & Schmidt (2002)	RS9	ME
Van Gerven, Paas, van Merriënboer, & Schmidt (2002a)	RS9	ME
Van Gerven, Paas, van Merriënboer, & Schmidt (2002b)	PR	
Van Gerven, Paas, van Merriënboer, & Schmidt (2002c)	RS9, ST	ME
van Merriënboer, Schuurman, de Croock, & Paas (2002)	RS9	ME

Note. Studies are listed in chronological order. PS = production system; ST = secondary task technique; RS = rating scale (9-point or 7-point scale); ME = mental efficiency; HRV = heart rate variability; PR = pupillary responses.

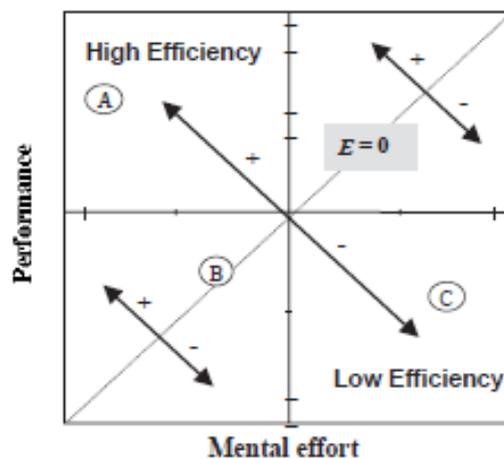


FIGURE 2 Graphic presentation used to visualize instructional efficiency.