

輪講2015/6/8

A Comparison of Three Measures of Cognitive Load:
Evidence for Separable Measures of Intrinsic, Extraneous,
and Germane Load

Krista E. DeLeeuw and Richard E. Mayer

発表者:水野陽介

- オセロ学習支援システム
- 学習への支援レベルを調整するのではなく、認知負荷を操作してはどうか？

認知負荷 (CLT, Cognitive Load Theory)

- 課題内在性負荷 Intrinsic Load
 - 課題の困難度が高いほどor課題に対する学習者の熟達度が低いほど、大きくなる
- 課題外在性負荷 Extraneous Load
 - 学習に必要なものもかかわらず生じる
 - 例: UIが悪い
 - 必要な関連知識が存在しない場合でも生じる
- 学習関連負荷 Germane Load
 - 学習のために使われる認知資源
 - 課題内在性負荷の一部であると考えられる
- 学習支援は学習者の処理能力を超えないように設計されなければならない
- マルチメディアを用いた学習中にどのように認知負荷を計測し、どのように操作するのかについては検討されている途中

3種類の認知負荷を操作する

認知負荷	認知負荷を操作する要因と例
課題外在性負荷	レッスン内容の冗長性 冗長性が高い例: アニメーション+ナレーション+字幕 →課題外在性負荷up 冗長性が低い例: アニメーション+ナレーション →課題外在性負荷down
課題内在性負荷	文章の複雑さ 複雑な例: 多くの相互に作用しあう概念 →課題内在性負荷up 複雑でない例: 少しの" →課題内在性負荷down
学習関連負荷	ポストテストにおける成績 良い成績: 学習中に得た知識を組み合わせさせて答えを作っている。 →学習関連負荷high 悪い成績: 学習中に得た知識を正しく組み合わせられていない。 →学習関連負荷low

本研究の目的

- よく使われる認知負荷計測の3つの指標(学習中のサブ課題への反応時間・学習中の努力指標・学習後の難易度評定)が3種類の認知負荷のどれに大きく影響するのかを調べる。

指標	課題の内容
学習中のサブ課題への反応時間	レッスン中に画面の背景がピンクから黒に変化し始めた時に、できるだけ速くスペースキーを押すというタスクを行う
学習中の努力指標	レッスン中に中断し「今のパートでの心的努力を1(とても低い)~9(とても高い)で表してください」と質問する
学習後の難易度評定	学習後に「どれだけこのレッスンが難しかったか」を1(とても簡単)~9(とても難しい)で答えさせる

実験1

- 目的: 課題外在性負荷・課題内在性負荷・学習関連負荷を操作しようと考案した3つの操作(サブ課題への反応時間・努力指標・難易度評定)が正しいかどうかを調べる
- 参加者: 56名の大学生
- 材料: 電気モーターがどのように動くかについて解説を行うプログラム

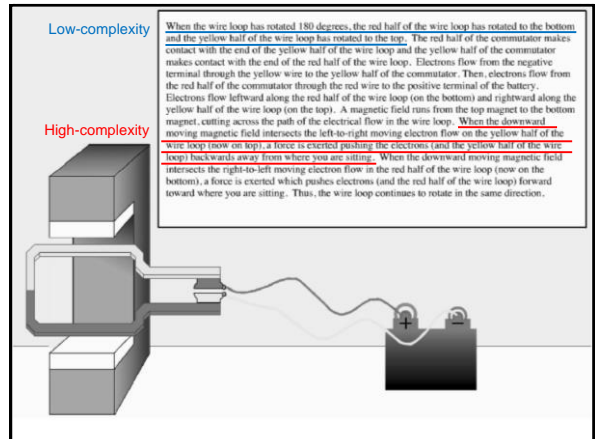
手続き

• プレテスト

- コンピューターに回答する形式
- 電気についての知識と経験について
- 年齢、性別、SATスコア、大学での学年

• マルチメディアレッスン(2回)

- 低冗長性プログラム群(アニメ+ナレーション) : 28名と高冗長性プログラム群(アニメ+ナレーション+字幕) : 28名
- どちらのプログラムにも複雑な解説文2つと複雑でない解説文2つが含まれる
- それぞれの文が流れたらサブ課題(画面の色が変化したらスペースバーを押す)と努力指標の回答を行う
- 努力指標の回答を終えると次の解説文へ進む(8回繰り返す)



手続き(続き)

• 難しさの評定&ポストテスト7問

• ポストテストの問題例

- モーターのスピードを上げるにはどうしたらよいか?
- スイッチを入れてもモーターは動かなかった。どこが悪いと考えられるか?

結果

Table 4
Means (and Standard Deviations) for Three Types of Cognitive Load Manipulations Based on Three Measures of Cognitive Load:
Experiment 1

Measure of cognitive load	Type of cognitive load					
	Extraneous load: Redundancy (Which cognitive load measure(s) is sensitive to redundancy?)		Intrinsic load: Complexity (Which cognitive load measure(s) is sensitive to sentence complexity?)		Germane load: Transfer (Which cognitive load measure(s) is sensitive to transfer performance?)	
	Redundant	Nonredundant	High	Low	High (n = 27)	Low (n = 27)
Response time (ms)	2.657 (825)	2.249 (719)	2.555 (1.035)	2.337 (714)	2.477 (933)	2.414 (636)
Effort rating	4.97 (1.54)	5.49 (1.41)	3.43 (1.35)	3.05 (1.50)	5.30 (1.57)	5.18 (1.42)
Difficulty rating	5.15 (1.54)	5.36 (1.47)			4.63 (1.39)	5.09 (1.34)

Note. N = 54. Boxes with bold lines indicate significant difference ($p < .05$). Boxes with light lines indicate $.05 < p < .10$ and effect size greater than $d = .20$.

- 2名はデータに不備があったので除外し、54名で分析
- ポストテストの成績は、回答をレッスンでの学習内容で得た知識を正しく使えているのかという基準で25点満点で採点。中央値より上の得点者をHigh群、下をLow群とした
- 学習関連負荷(germane)は検定、課題内在性負荷と課題外在性負荷は2×2の混合計画の分散分析を行った。
- 太枠が $p < .05$ で有意、細枠が $p < .10$ の有意傾向がみられた場所

結果

• サブ課題への反応時間について

- 冗長性高群・複雑性が高い文章の後・ポストテスト成績低群で反応時間は長くなるのか?
- しかし、有意差が見られたのは冗長性と複雑性のみ。
- サンプル数が少なかったからではないか?

• 努力指標について

- 冗長性高群・複雑性が高い文章の後・ポストテスト成績低群で努力指標は高くなるはず?
- しかし、複雑性のみ有意

• 難易度評定について

- ポストテスト成績において有意

考察

• この計測方法は別の認知負荷を計測しているのではないか?

- もしそうなら、相関関係があるはず
- しかし、相関関係は見られなかった

Table 5
Correlation Matrix of Dependent Measures for Both Groups:
Experiment 1

Measure	1	2	3	4
1. Dual-task reaction time	—	.27*	.20	.07
2. Self-report mental effort		—	.26	.19
3. Lesson difficulty rating			—	-.22
4. Score on transfer test				—

Note. N = 54.

* $p < .05$.

考察

- サブ課題への反応時間について
 - 冗長性高群・複雑性が高い文章の後・ポストテスト成績低群で反応時間は長くなるはず
 - しかし、有意差が見られたのは冗長性と複雑性のみ。
 - サンプル数がすくなかったからではないか？
- 努力指標について
 - 冗長性高群・複雑性が高い文章の後・ポストテスト成績低群で努力指標は高くなるはず
 - しかし、複雑性が高い文章の後のみ有意差が現れた
- 難易度評定について
 - ポストテスト成績において有意差が現れた。

実験2

- 実験1では、認知負荷を計測する方法が学習者にとって邪魔であり、認知負荷を高めるといってより気を逸らしてしまったのではないか？
- そこで、学習者のモチベーションを高め、より集中してもらうために、プレテストにおいてレッスンの後に質問される内容を教えた。
- 参加者：99名
- 実験デザインは実験1とほぼ同じ
- 実験1と異なる点はプレテストで、ポストテストに関する情報を提供すること
 - 「モーターのスピードを上げるにはどうしたらよいか？」等を2種類提供。「これはあとで聞かれる質問の例です」と教示

結果

Table 7
Means (and Standard Deviations) for Three Types of Cognitive Load Manipulations Based on Three Measures of Cognitive Load: Experiment 2

Measure of cognitive load	Type of cognitive load					
	Extraneous load: Redundancy (Which cognitive load measure(s) is sensitive to redundancy?)		Intrinsic load: Complexity (Which cognitive load measure(s) is sensitive to sentence complexity?)		Germane load: Transfer (Which cognitive load measure(s) is sensitive to transfer performance?)	
	Redundant	Nonredundant	High	Low	High (n = 45)	Low (n = 51)
Response time (ms)	2,918 (872)	2,520 (797)	2,677 (869)	2,769 (974)	2,569 (848)	2,859 (847)
Effort rating	3.67 (1.39)	4.99 (1.67)	3.47 (1.66)	3.21 (1.68)	3.58 (1.58)	3.13 (1.71)
Difficulty rating	3.33 (1.83)	3.21 (1.74)			4.82 (1.92)	3.67 (1.56)

Note: Boxes with bold lines indicate significant difference ($p < .05$).

- 3名はデータに不備があったので除外し、96名で分析
- ポストテストの成績は、回答をレッスンでの学習内容で得た知識を正しく使えているのかという基準で25点満点で採点。中央値より上の得点者をHigh群、下をLow群とした
- 学習関連負荷(germane)の差は検定、課題内在性負荷と課題外在性負荷は2×2の混合計画の分散分析を行った。
- 太枠で囲まれているのが $p < .05$ 、細い線で囲まれているのが $p < .10$ の有意傾向が現れた場所

結果

- サブ課題への反応時間について
 - 有意差が見られたのは冗長性のみ
- 努力指標について
 - 冗長性高群・複雑性が高い文章の後・ポストテスト成績低群で努力指標は高くなるはず
 - 実験1と異なり、冗長性と複雑性で有意
- 難易度評定について
 - ポストテスト成績で有意

考察

- この計測方法は別の認知負荷を計測しているのではないか？
 - 実験よりは数値が高かったものの、相関関係は見られなかった

Table 8
Correlations Between Dependent Measures for Both Groups: Experiment 2

Measure	1	2	3	4
1. Dual-task reaction time		.12	.13	-.30**
2. Self-reported mental effort	—	—	.33**	.11
3. Lesson difficulty rating			—	-.22*
4. Score on transfer test				—

Note. $N = 96$.
* $p < .05$. ** $p < .01$.

考察

- ある認知負荷を測る方法は、別の認知負荷に影響があるのではないか？
 - それぞれの測定方法の効果量を検討した
 - 認知負荷計測の方法(反応時間・心的努力・難易度評定)はそれぞれ別の認知負荷を計測することを確認した
 - 反応時間 → 冗長性=課題外在性負荷
 - 心的努力 → 複雑性=課題内在性負荷
 - 難易度評定 → ポストテスト成績 = 学習関連負荷

Table 6
Effect Sizes for Three Types of Cognitive Load Manipulation Created Based on Three Measures of Cognitive Load

Measure of cognitive load	Type of cognitive load					
	Extraneous load (redundant vs. nonredundant)		Intrinsic load (high vs. low complexity)		Germane load (high vs. low transfer)	
	Exp. 1	Exp. 2	Exp. 1	Exp. 2	Exp. 1	Exp. 2
Response time	.51*	.45*	.25*	ns	ns	ns
Effort rating	ns	.35*	.25*	.16*	ns	ns
Difficulty rating	ns	ns	ns	ns	.92**	.48*

Note. Cohen's d is the measure of effect size. Difficulty rating does not apply to intrinsic load. Exp. = Experiment.
* $p < .10$. ** $p < .05$. *** $p < .01$.

結論

- 二つの実験から、認知負荷を計測する方法はそれぞれ別の認知負荷を測定していることが確認できた。
- 理論的な示唆
 - もし認知負荷が異なる要素から成り立つのであれば、学習状況によって異なる操作が必要なのではないか
- 実験からの示唆
 - 認知負荷を計測する方法が一部の種類の認知負荷しか計測していない可能性を考えるべき

将来的な課題について

- 今後の実験でも学習者が知識を持っていない分野を選ぶべきである
- 認知負荷を操作する他の方法もどのように検討されるべきである
- 本実験では認知負荷の計測において学習中に不自然な介入を行ったので、より自然な学習環境を用意すべきである