

第2回エビデンスに基づく実践と政策セミナー
(第17回エビデンスに基づく教育研究会研究大会)

教育の効果

(Hattie, J. (2009). *Visible learning*. London: Routledge.)

について

国立教育政策研究所
総括研究官 やまもりこうよう 山森光陽
(教育心理学)

2018年6月30日

この内容は個人的見解であり
国立教育政策研究所の公式見解ではありません

本日の内容

1. *Visible learning* について

Visible learning の特徴と問題点

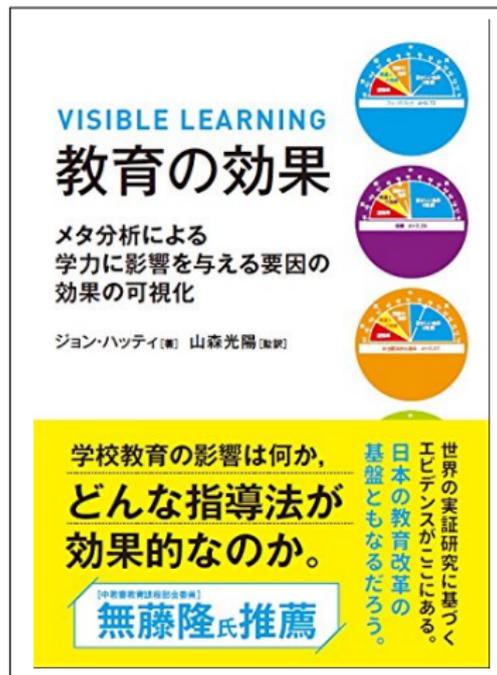
2. 平均値の高低と優越率

Common Language Effect Size または優越率

3. *Visible learning* の示唆

指導の効果と制度の効果に分けた議論の必要
更新される統合後効果量
効果の低い要因に対する解釈

「教育の効果」



- ▶ 学力に影響を与える諸要因を可能な限り取り上げ、各々に関する数多の実証研究で示された効果を統合し、学力に対する効果を網羅的に提示。
 - ▶ 800 本程度のメタ分析の結果を統合。この統合には 5 万本以上の一次研究。延べ対象者数は 8,800 万人超。
- ▶ 税込み 3,396 円だが、最近でも 1 日に 10 冊以上売り上げている。

和訳に至る経緯

「フィードバックは効果の大きい要因のトップ10に入る」

- ▶ 諸要因が学習成果に与える効果の順位表が存在？

最初の印象

- ▶ 自分の研究（特に学級規模→学力）で共変量にあたりをつけるのに便利。

ある日突然、翻訳著作権がとられ、監訳を振られる

- ▶ ある出版社の研究部門が読書会をしていた。
 - ▶ 「どうせなら訳して売れば？」
- ▶ 数年前で原著が3万部売り上げていて、様々な国に影響を与えていたにもかかわらず、なぜか著作権がとられていなかった。
 - ▶ 「著作権をとったので監訳してください」

Visible learning の特徴

スーパーシンセシスの結果の列挙

- ▶ 系統的レビュー（まとめ）の統合（まとめのまとめ）。
- ▶ 教育研究の中では「レベルの高いエビデンス」と言える。
- ▶ 学力に影響を与えうる諸要因（学習者、家庭、学校、教師、指導）を可能な限り取り上げている。
（原著は 138 要因，和訳は教育課程要因を中心に除外し 78 要因）

様々な要因を同一尺度で比較

- ▶ 効果量 d : SD を単位とした分布のずれ。
- ▶ 教育条件 < 教師による指導。

効果に対する説明

- ▶ 「児童生徒の能力を着実に伸ばす教師... の一連の特徴... の記述」

Visible learning 原著の問題点

引用, 数値, 計算の間違い

- ▶ 直接引用の文が引用元に存在しない箇所がある。
- ▶ 前後の内容と全く関係のない段落がある。
- ▶ 効果の背景を説明するために用いた理論の記述を間違っている箇所がある。
- ▶ r と d を間違えている箇所がある。 r から換算した d の値が違う箇所もある。

分類の適切さ

- ▶ 学年や年齢の考慮なしに二次研究を分類カテゴリに投入していると思われる部分がある。
- ▶ 分類カテゴリが大括りだったり細に入ったりしていると思われる部分がある。

CLE が全て間違っている

	<i>d</i>	SE	CLE
<i>Background</i>			
Prior achievement	0.67	0.098	48%
Piagetian programs	1.28	—	91%
Self-report grades	1.44	0.030	102%
Creativity	0.35	—	25%
<i>Attitudes and dispositions</i>			
Personality	0.19	0.007	14%
Self-concept	0.43	0.010	30%
Motivation	0.48	0.047	34%
Concentration/ persistence/engagement	0.48	0.032	34%
Reducing anxiety	0.40	—	28%
Attitude to mathematics/science	0.36	—	26%
<i>Physical influences</i>			
Pre-term birth weight	0.54	—	14%
Illness	0.23	—	16%

(a) 原著 Table 4.1 (抜粋)

学習者要因	効果量 (<i>d</i>)	標準 誤差	CLE
<i>学習者の経歴</i>			
過去の学力	0.67	0.098	68%
ピアジェによる発達段階*	1.28	—	81%
能力レベルの自己評価*	1.42	0.030	80%
創造性*	0.35	0.110	60%
<i>態度と気質</i>			
性格	0.19	0.007	55%
自己概念	0.43	0.010	62%
動機づけ	0.48	0.047	63%
集中力, 我慢強さ, 積極性	0.48	0.032	63%
不安の低減*	0.40	—	61%
教科に対する態度	0.36	—	60%
出生時の体重*	0.54	—	65%
病気*	0.23	—	56%

(b) 訳本 表 4.1 (抜粋)

Figure 1: CLE の違い

「内容の半分は間違い」？



The screenshot shows a web browser displaying the Save Our Schools NZ website. The URL in the address bar is <https://saveourschoolsnz.com/2014/09/26/john-hattie-admits-that-half-of-the-s...>. The page features the Save Our Schools NZ logo and tagline: "FREE, FAIR, EQUITABLE PUBLIC EDUCATION FOR ALL". A navigation menu includes "HOME", "MISSION STATEMENT", "NZ CHARTER SCHOOLS - VIDEO", and "MAKE A CHANGE". A red banner with a tomato graphic contains the text "you're reading... Save Our Schools NZ". Below the banner, the article title is "John Hattie admits that half of the Statistics in Visible Learning are wrong". The post is dated "SEPTEMBER 26, 2014" and has "1 COMMENT". It is categorized under "FILED UNDER CLE, ERRORS IN STATISTICS, HATTIE, JOHN HATTIE, , VISIBLE LEARNING". The article text begins: "Given Hattie's influence in the education arena, particularly in policy making, he really should be far more careful to ensure his calculations and analysis is correct. This is abysmal, really." A link is provided: "See also: <http://academiccomputing.wordpress.com/2013/08/05/book-review-visible-learning/>".

Figure 2: Hattie が統計量の半分が間違いであることを認める

CLE に対する Hattie の釈明

research **ED**

Home About News Events Magazine Resources Contact us

Home > Magazine > John_Hattie_interview

John Hattie on Visible Learning

Interviewed by the employee magazine educ.alla Utbildningsförvaltningen - Göteborg



Visible Learning has had a major impact in Sweden and among your results the study shows a very small effect of smaller groups or classes and of homework assignments in certain age groups. Can you see a change in attitude around these issues among school professionals and connecting groups, or are there still many who maintain that these factors have a large significance for learning? Is there a difference between groups (parents, students, teachers, politicians)?

The key issue with influences that have very small effects (and many that have effects much smaller than many would expect or hope for) is to ask "Why are they so small?" The reason, I would argue, that smaller class sizes make little difference is because teachers have not changed their teaching methods as they move from larger to smaller. The reason any homework or elementary schools is too is that the homework is often a "task" (or a project) and not an opportunity to reinforce the learning that has begun in the classroom.

It is certainly the case that many do not want to believe evidence as their own "experiences" tell them different. Research starts from the premise of attempting to find benefits and evidence to support their prior beliefs. This confirmation bias means they spend millions of dollars on the wrong issues, and thus do major damage to the learning of our children.

Your study of meta-analysis is very comprehensive, but I understand you keep adding to it. How does the database look today? Have you discovered new factors that have significant impact on learning and by that concern the learning profession?

Yes I am now close to 1200 meta-analyses (up from 800 in V1). What is remarkable since I published the first study (in 1998 based on 134) is that the "story" underlying the data has hardly changed. Some of the more interesting (new) effects include DuBryan (-.24), parental employment (.03), sleep (.07), single sex compared to coed schools (.08), Philosophy in schools (.43), Service learning (.09), conceptual change programs (.10), and collective self-efficacy (.10).

As always, when it comes to research, there are some objections to your results. What would you yourself say is the most controversial area of your research?

Meta-analysis is a well-rehearsed and well known method. While some bloggers have trouble with effect-sizes, they are an established part of the statistical scene. Other bloggers claim that even I say that "half the data in V1 is wrong" - how absurd and I have never made such a claim! Yes, I made a mistake calculating on statistics (the CLE) in early editions of V1, but it is hardly important given the attention to the effect-sizes (although I regret that error) and the very minor part played by the CLE (which almost every reader and critic has ignored anyway). I would have wanted more moderators in the literature other than the more obvious and often less interesting (the achievement levels compared to gender, age) - but this is a reflection on the research and methods. Also, too many studies are based on surface rather than deep learning, but again this is a reflection of what so often occurs in our classrooms. The data from the meta-analysis is hardly controversial but the more important part is the "explanation" - I have got to find any alternative story from the one I proposed in the various V1 books.

Figure 3: Hattie による釈明

Hattie の釈明

- ▶ 計算間違えました。
- ▶ (CLE と d を含めて) 効果量の指標に着目することが重要なのではない。
- ▶ CLE は脇役に過ぎない (の でほぼ全ての読者は着目していない)。

その後

- ▶ 続編のような書籍では CLE には触れないようになって いる。

CLE または優越率

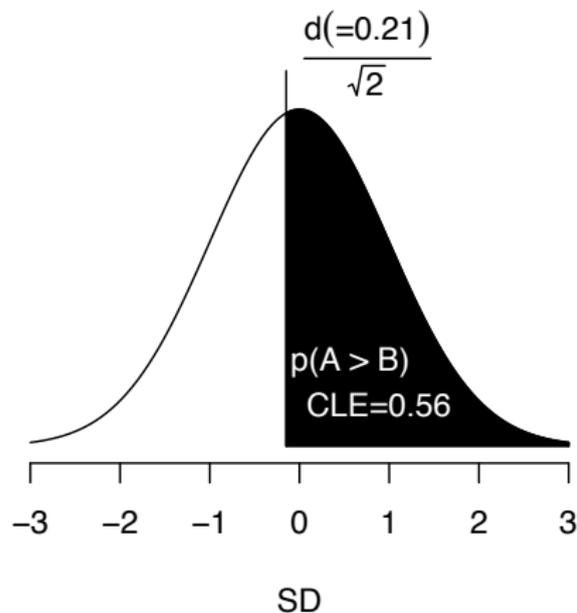


Figure 4: CLE または優越率

- ▶ 一方の群のケースが他方の群のケースを上回る確率。
 - ▶ CLE[McGraw and Wong, 1996]
 - ▶ 優越率: 南風原・芝 (1987)
- ▶ 優越率: $\pi_d = \Phi(d/\sqrt{2})$ で求められた, 標準正規分布の下側確率 (対応ありでは $\Phi(d)$)。
- ▶ *Visible learning* での学級規模 $d = 0.21$ の場合 $\pi_d = 0.56$ [Hattie, 2009]

優越率の解釈

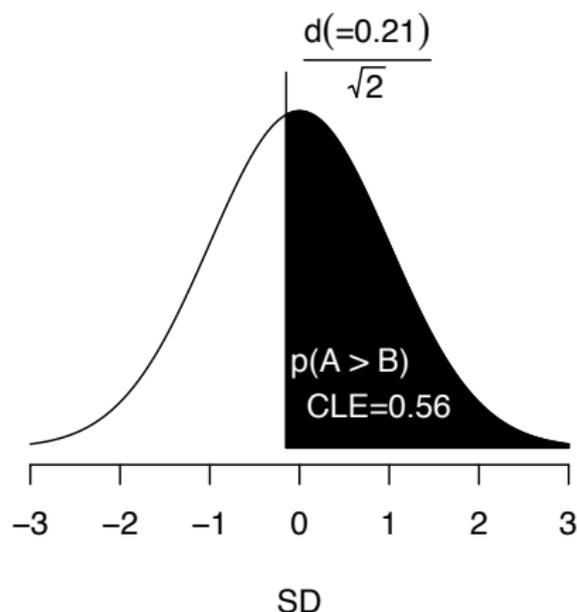


Figure 5: CLE または優越率（再掲）

学級規模 ($d = 0.21$) の場合

- ▶ 小規模学級にアサインされた 40 人の児童のうち 22 人が大規模学級にアサインされた児童の得点を上回る。
- ▶ 大規模学級にアサインされた 40 人の児童のうち 18 人が小規模学級にアサインされた児童の得点を上回る。

Visible learning [Hattie, 2009] での CLE はスーパーシンセシスの対象となった二次研究ごとに求めた CLE の平均。

効果量 d と優越率の関係

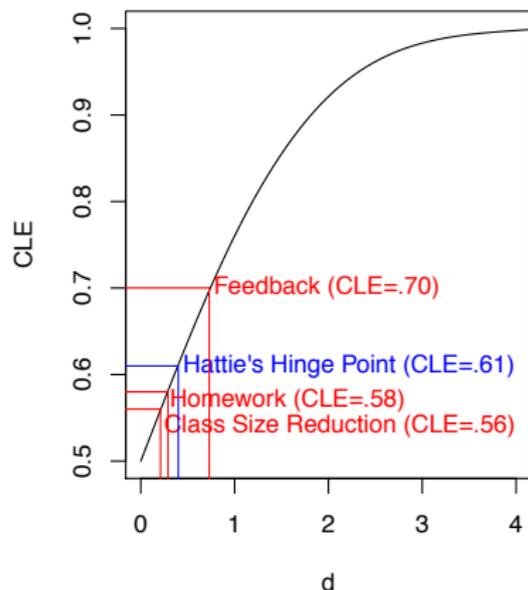


Figure 6: 効果量 d と優越率の関係

- ▶ Hattie が Hinge point (効果の基準値) としている $d = 0.40$ でも優越率は 0.61。
- ▶ $d = 1.00$ で優越率は 0.76。
- ▶ 平均値差の大きさだけでなく、「どの程度の児童生徒が得をするか」といったことを示す指標を解釈する必要もあるのではないか。
- ▶ 「誰に効果的か」をイメージしやすい可能性。

指導の効果と制度の効果

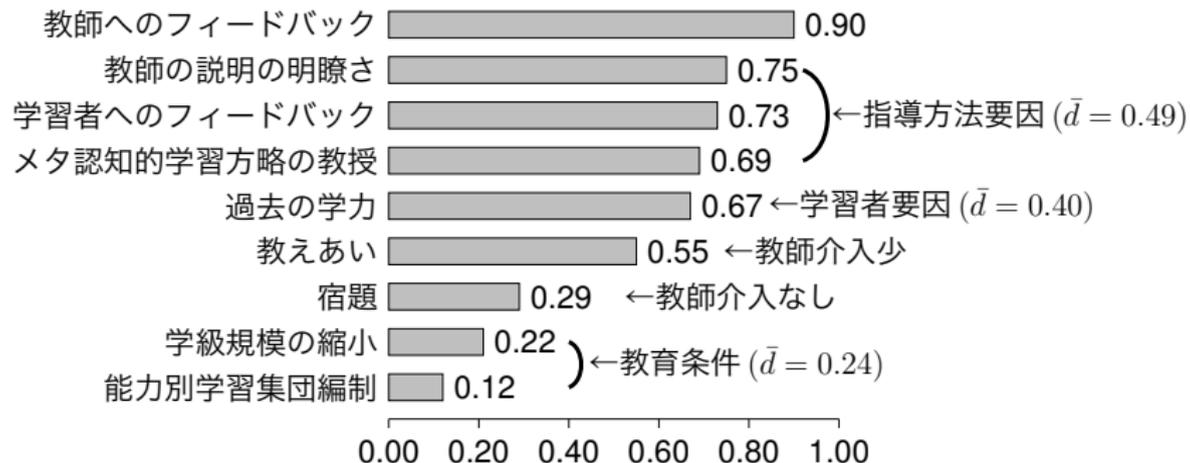


Figure 7: 指導の効果と制度の効果の比較

教育条件と指導を同列に扱った議論

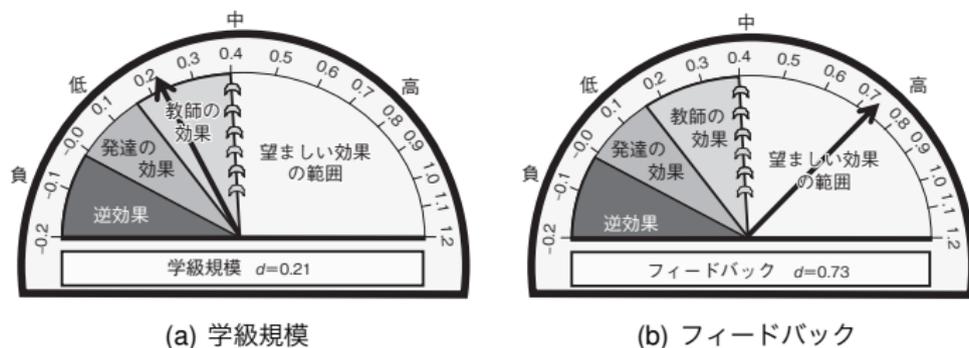
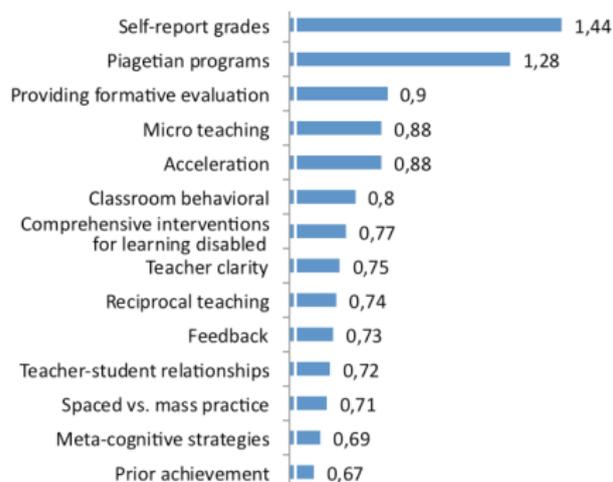


Figure 8: 学級規模とフィードバックの効果 [Hattie, 2009]

- ▶ $0.3SD$ 分の学力の向上を目指す介入を行うと仮定
- ▶ 誤りなどに対する即時フィードバックで $0.3SD$ 分の学力の向上につなげることを教師が出来るようにするための研修のコストは子ども一人当たり 28.31 ドル (約 3,000 円)。
- ▶ 学級規模の縮小で $0.3SD$ 分の学力の向上につなげることを教師が出来るようにするための研修のコストは子ども一人当たり 1,379.28 ドル (約 150,000 円) [Yeh, 2009]。
- ▶ 学級規模が大きくても効果的なフィードバックは可能か。

効果は固定的なものではない



(a) 2009



(b) 2018

Figure 9: Hattie Ranking (抜粋) <https://visible-learning.org/>

- ▶ Feedback: 0.73 → 0.70 など数値に違い。
- ▶ Jigsaw, Discussion, Deliberate practice, Help seeking など追加。

効果の低い要因に対する解釈

d がほぼゼロ

- ▶ 例：オープン教育 ($d = 0.01$)，異年齢集団編制 ($d = 0.09$)
- ▶ オープンスクールであるかないか，複式学級か単式学級かといった違いそのものは学力に影響しない。
- ▶ 所与の条件で実施しやすい効果的な指導を実施することでアウトカムの向上につながる。

d が高くない教育条件

- ▶ 例：学級規模 ($d = 0.21$)，能力別学習集団編制 ($d = 0.12$)
- ▶ 優越率も考慮して（学級規模の場合 0.56）誰に効果的か，他の子どもに効果的となるようにするにはどのような介入を行えばよいかを検討する必要。

思いの外 d が低い介入

- ▶ 例：宿題 ($d = 0.29$)
- ▶ 効果のある子とない子がいる。どうすれば思っているような効果となるか別の要因の効果を参照しながら介入の仕方を変化させる。

- ▶ 効果が低いことが「実施すべきではない」ことを意味するのではない（正負で言えば正）。

Visible learning の示唆

あらゆる要因の効果の同一尺度上での表現

- ▶ 学習者、家庭、学校、教師、指導の各類型に対して様々な要因、介入の効果の比較可能化。
- ▶ 各類型の効果の範囲を示すことで、類型をまたいだ議論に慎重さが必要なことを示唆。

メタ分析は過去のこと、エビデンスは可変的

- ▶ 「メタ分析が基づいているのは『過去』の研究であり、昨日上手くいったからといって未来はそれに縛られるようなものではない」
- ▶ 新規の一次研究 → これらを含めた二次研究 → これらのスーパーシンセシス → 「エビデンス」は変わる。

教室は特定の単独の介入「だけ」が行われる場所ではない

- ▶ 「他の変数との相互作用が結論を劇的に変えてしまうこともある」

まとめ

Visible learning について

- ▶ 内容に問題のある箇所もあるが，邦訳では可能な限り修正。
- ▶ 英，独，諾，丁，瑞，中の各国語版があるが，邦訳は「世界で一番正しい *Visible learning*」

参照すべき効果の指標

- ▶ 原著者 Hattie が「脇役」と呼んだ CLE（優越率）も参照する必要。

Visible learning で示された「効果」

- ▶ 学力に関わる諸要因に対する網羅的なスーパーシンセシス。
- ▶ 過去のこと，変わるもの，変えられるもの。

まとめ

他の変数との交互作用

- ▶ 「効果は変動しうるものである… 効果量… は、すべての学習者、すべての要因にあてはまるものとは限らない。効果量を左右する要因も多々ある」
- ▶ 「指導方法の組み合わせによっては、かけ算的な効果が期待」

網羅的な「エビデンス」とのつきあい方

- ▶ 「エビデンスとは、行動の指針を示すものではなく、知的な問題解決における仮説、あるいは教育の最終目標とは何たるかを探究する際の仮説をもたらすものに過ぎない」 [Dewey, 1938]

- ▶ 機械的な選択ではなく、問いを持つこと。

引用文献（欧文のみ）

[Dewey, 1938] Dewey, J. (1938).

Logic : the theory of inquiry.

Holt, Rinehart and Winston, New York.

[Hattie, 2009] Hattie, J. (2009).

Visible learning: A synthesis of over 800 meta-analyses relating to achievement.

Routledge, London.

[McGraw and Wong, 1992] McGraw, K. O. and Wong, S. P. (1992).

A common language effect size statistic.

Psychological Bulletin, 111:361–365.

[Yeh, 2009] Yeh, S. S. (2009).

Class size reduction or rapid formative assessment? a comparison of cost-effectiveness.

Educational Research Review, 4:7–15.