Instructional Interventions Affecting Critical Thinking Skills and Dispositions: A Stage One Meta-Analysis of Empirical Evidence

E. Borokhovski, P. C. Abrami, R. M. Bernard, A. Wade, M. A. Surkes, R. Tamim, E. Bethel, T. Persson, & D. Zhang

Centre for the Study of Learning and Performance

Concordia University, Montreal, Quebec CANADA

Campbell Collaboration annual meeting, Vancouver, May 14, 2008

Interests in CT

- Longstanding interests in critical thinking

  Critical thinking (CT), or the ability to engage in purposeful, self-regulatory judgment, is widely recognized as important, even critical aspect of any educational practice, by governments, educators, and employers, both nationally and internationally and has a long history (e.g., Watson, 1942)

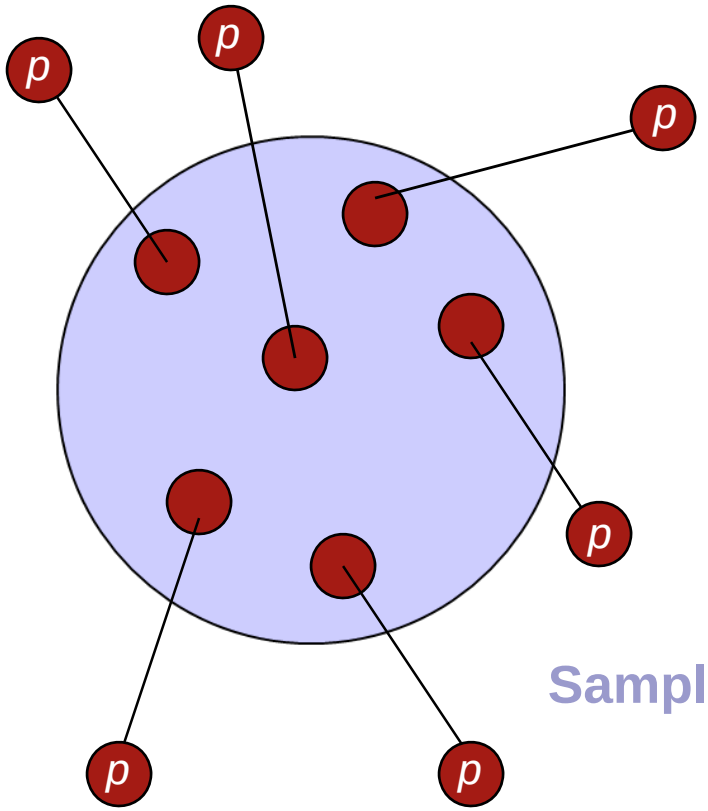
- This meta-analysis conducted at CSLP at Concordia University (Montreal, Canada) investigates the link between instructional interventions and CT outcomes

Research question and method:

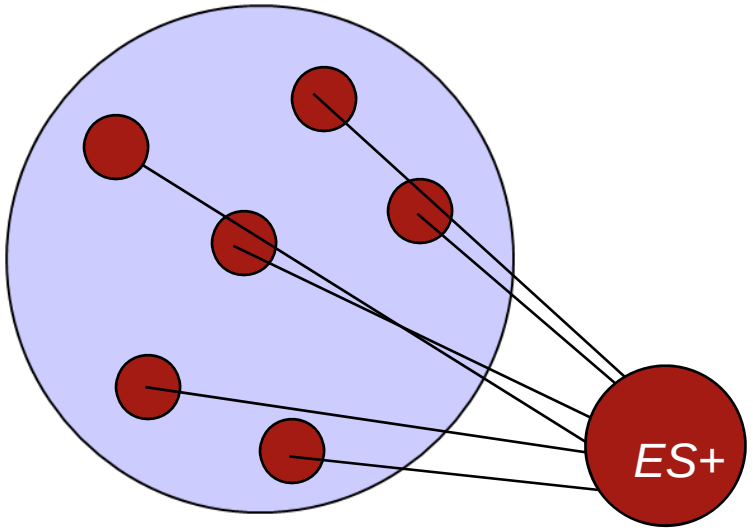
- **What instructional interventions, to what extent, and under what circumstances, have an affect on the development and effective use of CT skills and dispositions?**

- We have chosen the method of a meta-analysis because it not only shows whether a general effect exists and what its overall magnitude is, but also allows for exploring sources of variability in effect sizes through an analysis of methodological and substantive study features as they mediate and/or moderate the influences of experimental treatments on CT.

What is Meta-Analysis



Samples:



Meta-Analysis:

Instead of the outline:
Some potentially controversial issues

- Defining CT
- **Is critical thinking a generic skill & what instructional methods are better suited to teach it?**
- What is known about teaching students to think critically?
- Outcome measures
- **Research design & other aspects of methodological quality**
- Age of learners
- Treatment duration
- **Categorizing treatments**

Issues addressed to enable the full-scale meta-analysis:

- Is critical thinking a generic skill & what instructional methods are better suited to teach it?

  CT typology of four courses — *Generic, Infusion, Immersion, and Mixed* — Ennis (1989)

  Theoretical frameworks / Pedagogical models (approaches)

Issues addressed to enable the full-scale meta-analysis:

- Research design & other aspects of methodological quality

  Inclusive approach / Coding for methodological study features

  Statistical control procedures (Abrami & Bernard, under review)

  External / Internal validity

  Construct validity (Treatment / Outcome measures)

Issues addressed to enable the full-scale meta-analysis:

- Research design & other aspects of methodological quality

Treatment "Fidelity": Ascending from 1 to 3*
CT measure: 3 = Standard, 2 = Derivative, 1 = Theoretical
Cross-tabulation Count

		CT measure			Total
		1	2	3	
Treatment "Fidelity":	1	0	0	3	3
	2	27	5	37	69
	3	10	6	13	29
Total		37	11	53	101

Spearman Correlation: - .085, p = .396,
Kendall's tau: - .080, p = .380

Treatment duration: Ascending from 1 to 4 *
Qualty of research design: Ascending from 1 to 3
Cross-tabulation Count

		Qualty of research design			Total
		1	2	3	
Treatment duration:	1	1	7	5	13
	2	5	15	8	28
	3	20	36	4	60
	4	9	35	10	54
Total		35	93	27	155

Spearman Correlation: - .066, p = .417
Kendall's tau: - .056, p = .454

Issues addressed to enable the full-scale meta-analysis:

- Categorizing treatments:

  Coding procedure
- Emerging Study Features:

  1. Feedback (including graded CT related assignments)

  2. “Real life” applicability (linking treatment to future job and/or authentic cases/problems to study/solve)

  3. Explicitly addressed any of 6 CT major skills

  4. Use of technology

  5. Use of a sound model/framework to guide instruction

  6. The emphasis on interdisciplinary connections

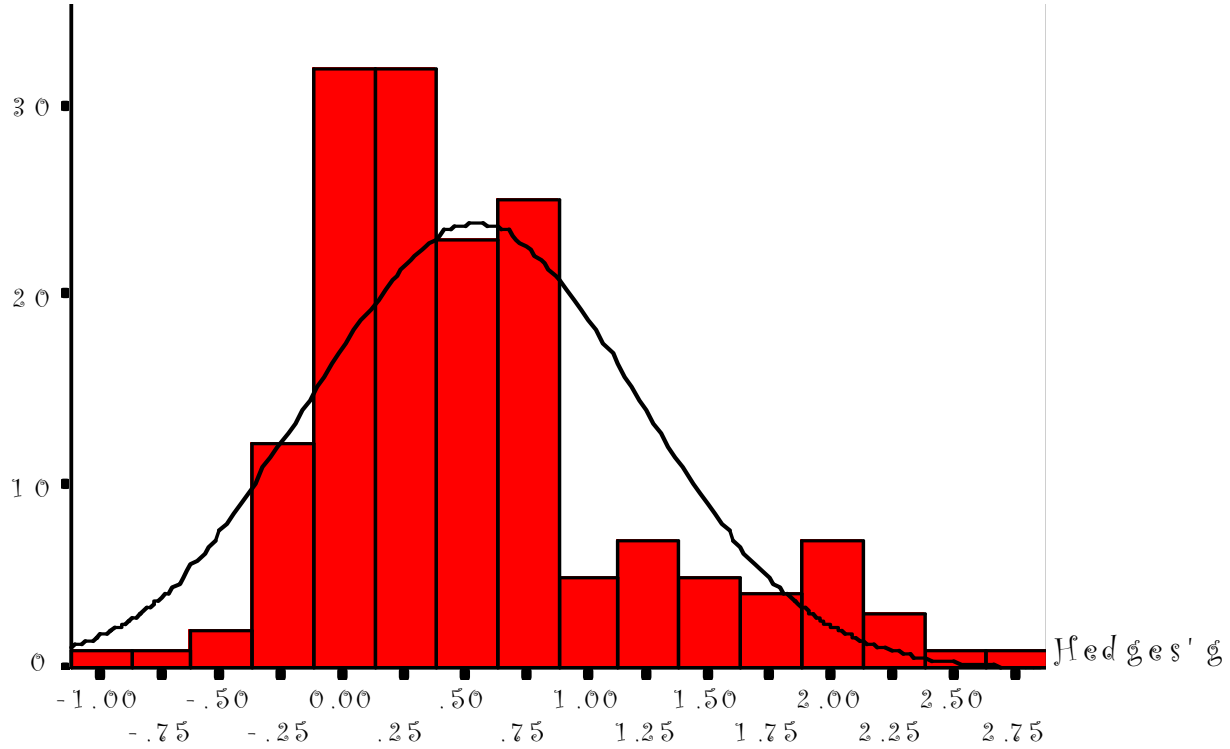
  7. Instruction in logic etc.

  8. Meta-cognitive self-regulation
- Coding for “treatment fidelity”

Methodology:

- Literature search strategies
  - Key terms
  - Data bases
- Inclusion/Exclusion criteria
  - Accessibility, Relevance, Intervention, Age, Research Design, etc.
- Effect size extraction
  - Cohen's *d* & Hedges' *g*
    (Glass, McGaw, & Smith, 1981, Hedges & Olkin, 1985)
- Study features coding and analyses
  - Research design, CT measure used, type of intervention, pedagogical relevance, students' age, duration of treatment, etc.

Major results:



Distribution of effect sizes

Major results:

Table 1a. Overall mean effect size ($g+$) and statistics.

Overall Outcome	k	Effect size		95% confidence interval		Test of null
		$g+$	SE	Lower	Upper	z
Summary Statistics	161	0.34	0.01	0.31	0.37	23.58*

*$p < 0.001$

Table 1b. Heterogeneity analysis (Q and I^2) for overall mean effect size.

Overall Outcome	Heterogeneity			
	Q-value	df (Q)	p	I^2
Summary Statistics	1767.86	160	0.00	90.95

Major results:

Table 3a. Effect size ($g+$) for each type of research design.

Research Design	k	Effect size		95% confidence interval		Test of null
		$g+$	SE	Lower	Upper	z
1. Pre-Experimental	60	0.31	0.02	0.26	0.36	12.73*
2. Quasi-Experimental	74	0.36	0.02	0.32	0.40	18.01*
3. True-Experimental	27	0.34	0.04	0.26	0.42	8.49*

* $p < 0.01$

Table 3b. Heterogeneity analysis (Q and I^2) for type of research design.

Research Design	Heterogeneity			
	Q-value	df (Q)	P	I^2
1. Pre-Experimental	406.77	53	0.00	94.464
2. Quasi-Experimental	1099.67	86	0.00	86.796
3. True-Experimental	259.14	26	0.00	85.659
Total Within (Q_W)	1765.58	158	0.00	
Total Between (Q_B)	2.28	2	0.32	
TOTAL (Q_T)	1767.86	160	0.00	90.95

Major results:

Table 4a. Effect size ($g+$) for type of measure.

Type of Measure	k	Effect size		95% confidence interval		Test of null
		$g+$	SE	Lower	Upper	z
1. Standardized	91	0.24	0.02	0.20	0.28	12.29*
2. Teacher-made	7	1.43	0.12	1.19	1.66	11.78*
3. Researcher-made	31	0.65	0.03	0.59	0.71	21.23*
4. Teacher/Researcher	17	0.16	0.04	0.08	0.23	4.27*
5. Secondary-source	15	0.28	0.05	0.18	0.39	5.44*

* $p < .01$

Table 4b. Heterogeneity analysis (Q and I^2) for type of measure.

Type of Measure	Heterogeneity			
	Q-value	$df(Q)$	p	I^2
1. Standardized	566.10	90	0.00	84.10
2. Teacher Made	31.92	6	0.00	81.20
3. Researcher Made	654.55	30	0.00	95.42
4. Teacher/Researcher	199.02	16	0.00	91.96
5. Secondary-source	83.99	14	0.00	83.33
Total Within (Q_W)	1535.59	156	0.00	
Total Between (Q_B)	232.27	4	0.00	
TOTAL (Q_T)	1767.86	160	0.00	90.95

Major results:

Table 7a. Effect size ($g+$) for type of intervention.

Type of Intervention	K	Effect size		95% confidence interval		Test of null
		$g+$	SE	Lower	Upper	z
1. General CT Skills						
2. Infusion	52	0.54	0.03	0.49	0.59	21.14*
3. Immersion	48	0.09	0.02	0.05	0.13	4.10*
4. Mixed	22	0.94	0.06	0.82	1.05	16.16*

* $p < 0.01$

Table 7b. Heterogeneity analysis (Q and I^2) for type of intervention.

Type of Intervention	Heterogeneity			
	Q-value	$df(Q)$	p	I^2
1. General CT Skills	172.09	38	0.00	77.92
2. Infusion	939.46	51	0.00	94.57
3. Immersion	164.68	47	0.00	71.46
4. Mixed	195.78	21	0.00	89.27
Total Within (Q_W)	1472.02	157	0.00	
Total Between (Q_B)	295.84	3	0.00	
TOTAL (Q_T)	1767.86	160	0.00	90.95

Major results:

Table 8 a. Effect size ($g+$) for pedagogical grounding of intervention.

Pedagogical grounding of Intervention	k	Effect size		95% confidence interval		Test of null
		$g+$	SE	Lower	Upper	z
1. Instructor training	16	1.00	0.04	0.92	1.07	25.09*
2. Extensive Observations	46	0.58	0.04	0.51	0.65	16.62*
3. Detailed curriculum description						
4. CT among course objectives						

* $p < 0.01$

Table 8b. Heterogeneity analysis (Q and I^2) for pedagogical grounding of intervention.

Pedagogical grounding of intervention	Heterogeneity			
	Q-value	$df(Q)$	p	I^2
1. Instructor received training	532.43	15	0.00	97.18
2. Extensive Observations	295.53	45	0.00	84.77
3. Detailed curriculum description				
4. CT among course objectives				
Total within (Q_W)				
Total Between (Q_B)	446.16	3	0.00	
TOTAL (Q_T)	1767.86	160	0.00	90.95

Major results:

Table 9a. Effect size ($g+$) for the presence and absence of student collaboration in CT interventions.

Collaboration	k	Effect size		95% confidence interval		Test of null
		$g+$	SE	Lower	Upper	z
No	102	0.31	0.02	0.28	0.35	18.39
Yes	59	0.41	0.03	0.36	0.46	15.02

Table 9b. Heterogeneity analysis (Q and I^2) for the presence and absence of student collaboration in CT interventions.

Collaboration	Heterogeneity			
	Q-value	df (Q)	p	I^2
No	1031.95	101	0.00	90.31
Yes	726.65	58	0.00	92.02
Total within (Q_W)	1758.60	158	0.00	
Total Between (Q_B)	9.25	2	0.01	
TOTAL (Q_T)	1767.86	160	0.00	90.95

Discussion

- **The overall average effect size (g+) = 0.341 (SD = 0.610) varying dependent on SF (117 studies yielding 161 effect sizes, N=20698).**
- **All of the average effect sizes are significantly heterogeneous, suggesting that none can be considered as representing a population parameter.**
- **Type of CT intervention and pedagogical grounding were substantially related to fluctuations in CT effects sizes, together accounting for 32% of the variance.**
- **Improvement in students' CT skills and dispositions cannot be a matter of implicit expectation, educators must take steps to make CT objectives explicit in courses and also to include them in faculty development and training.**

Discussion
Your feedback on issues of:

- Defining CT
- Is critical thinking a generic skill & what instructional methods are better suited to teach it?
- What is known about teaching students to think critically?
- Outcome measures
- Research design
- Age of learners
- Treatment duration
- Categorizing treatments

THANK YOU!

- **We welcome comments and criticisms which may lead us to improve the quality of our undertaking**

- **Contact:** *Eugene Borokhovski*
  eborokhovski@education.concordia.ca

Concordia UNIVERSITY

CSLP/CEAP
CENTRE FOR THE STUDY OF LEARNING AND PERFORMANCE
CENTRE D'ÉTUDES SUR L'APPRENTISSAGE ET LA PERFORMANCE

Issues addressed to enable the full-scale meta-analysis:

- What is critical thinking? How is CT different from other thinking skills?

  CT as a composite of 6 skills and 19 dispositions - according to APA Delphi panel of experts (Facione, 1990)

  Problem-solving

  Creative thinking

  Meta-cognitive self-regulation

Issues addressed to enable the full-scale meta-analysis:

- What is known about teaching students to think critically?

  Inconclusiveness of previous reviews
  (e.g., Norris, 1985, McMillan, 1987)
  General maturation account
  CT dispositions & psychological personality traits

Issues addressed to enable the full-scale meta-analysis:

- How is CT measured?

  13 Major CT skill standardized tests and a variety of non-standardized measures

  Factor analysis of WGCTA (Bernard et al., AERA 2006)

Issues addressed to enable the full-scale meta-analysis:

- Age of learners

  “Formal operations”

  “Proximal development”, cognitive apprenticeship

  Vast variety of ages in empirical literature (e.g., “critical reading” in elementary school)

Issues addressed to enable the full-scale meta-analysis:

- Treatment duration

  Short duration studies are among the most carefully designed and controlled

  No artificial restriction of the scope of possible pedagogical approaches in CT instructional interventions

  3 hours of class time devoted to CT instruction should not be dismissed as irrelevant (especially when the subject matter and course objectives are not CT per se)