

COST-EFFECTIVENESS ANALYSIS FOR EDUCATION AND SOCIAL POLICY: THEORY AND PRACTICE

Clive R. Belfield

Economics Dept., City University of New York
Center for Benefit-Cost Studies in Education,
Teachers College, Columbia University

www.cbcse.org

Outline

Cost-effectiveness analysis (CEA):

Evaluation method that compares policy alternatives based on ratio of their costs to a quantifiable (but not monetized) effectiveness measure

- ① Current research practice in education
- ② CEA problems
- ③ Future for CEA

Almost no CEAs in education

- CEA is not a newly developed research method
 - Thousands of CEAs in health research
 - CEAs in education are becoming more common (with cost disease pressures and move to experimental methods)
 - But very slowly and from a low base; much CEA research is 'rhetorical'
- Policy reform debate is misguided and distorted

Misguided reform debate in education

- Reducing class size – effective but expensive
- Vouchers – supposedly ‘free’
- School-wide reforms – hidden reorganization costs
- Early reading programs – very different sizes
- Teacher accountability – wage effects ignored
- Web-learning, MOOCs – claimed ‘low-cost’ via large scale

- **Need economic analysis: CEA, CBA or other**

Center for Benefit-Cost Studies in Education

Teachers College, Columbia University

www.cbcse.org

- Perform CBA/CEA:
 1. Dropout prevention (2012)
 2. Early literacy (2013)
 3. Socio-emotional learning interventions (2014)
- Train researchers in CBA/CEA
- Provide toolkit (CEA spreadsheets, input price spreadsheets, inflation indices, locality indices, amortization calculator)
- Guidance/recommendations on best practice

CEA Problems: Costs

- ◆ Empirical problems with costs data:
 - Cost data reliant on budget documents
 - Marginal costs typically unavailable
 - Control group costs ignored
 - Data collected *ex post*
- ◆ Methodological problems with costs analysis:
 - Limited sensitivity testing
 - Limited statistical testing
 - Not harmonized across studies

CEA Problems: Validity

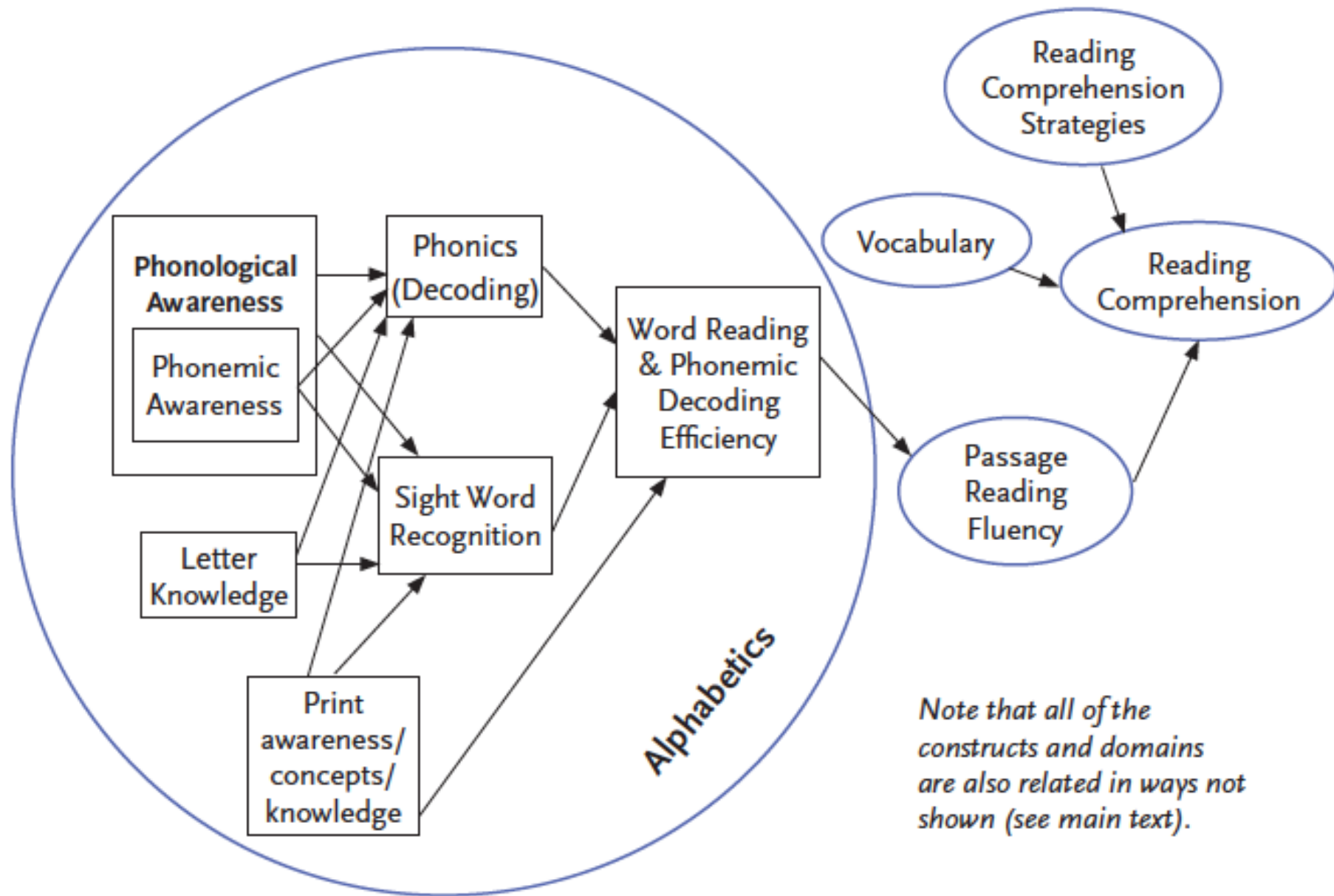
- CEA requires precise description of inputs of intervention
But many interventions lack fidelity with respect to site, duration, in-kind resources, or scale:
 - Off-shelf ‘standard’ reading programs per student cost \$400 to \$1,200
 - Talent Search delivery dosage of 1-6 years
- CEA forces precise description of *incremental* inputs of intervention relative to business as usual
But this – buying ‘gains’ in outcomes, not absolute outcomes – is hard to explain or to value
- Many interventions fail to specify ‘production function’ or ‘technology of skill formation’

**Site-level Cost-effectiveness Results for JOBSTART
(Social Costs; High School Graduation)**

	Cost per Participant	Gain in HSG (% point)	C-E Ratio (\$Cost/HSG)	Yield (Extra HSG per \$100,000)
All sites	\$10,460	15.1	\$69,510	1.44
By site:				
Corpus Christi	\$4,340	28.0*	\$15,520	6.44
El Centro	\$10,790	39.0*	\$27,670	3.61
Connelley	\$10,660	28.2*	\$37,810	2.64
Phoenix	\$12,480	20.7*	\$60,280	1.66
EGOS Denver	\$4,570	7.4	\$61,310	1.63
Allentown	\$11,920	9.5	\$125,150	0.80
Hartford	\$11,820	7.8	\$152,210	0.66
Atlanta	\$11,660	5.9	\$197,800	0.51
LA Jobs Corps	\$15,720	7.7	\$204,470	0.49
CET/San Jose	\$6,460	3.1	\$206,830	0.48
East LA	\$12,060	0.0	n/a	0.00
BSA (NYC)	\$20,190	-1.2	n/a	-0.06
Chicago	\$14,330	-5.0	n/a	-0.35

CEA Problems: Effectiveness

- ◆ Methodological problems:
 - Effects for follow-up subsamples with attrition – matters for costs
 - Effects reported on ITT or TOT – matters for costs
 - Effect size interpretation depends on variation within samples
- ◆ Validity problem: How to measure effectiveness? CEA forces a uni-dimensional answer
 - Unclear, multiple outcomes
 - Cumulative nature of learning
 - Cost-utility analysis not used (unlike QALYs)



Note that all of the constructs and domains are also related in ways not shown (see main text).

CEA Problems: Application

- ◆ CEA evidence implies decisions:
 - Not enough comparative evidence
 - Results depend on which decision-maker specified
 - Policymakers do not like this implication
 - Squeamishness of allocating resources
 - Decisions are too easy: peer tutoring (cheap student labor); targeted interventions (big effects); higher ability students (easier to reach thresholds)
 - Not the purpose of education research (Ludwig, Kling, and Mullainathan, 2011)

Cost-Effectiveness Ratios across Interventions to
Raise the High School (HS) Completion Rate

	Cost per Student	Cost per Extra HS Completer	Yield: Extra HS Completers per \$100,000
NGYC	\$14,100	\$71,370	1.40
Job Corps	\$22,290	\$131,140	0.76
JOBSTART	\$10,460	\$69,510	1.44
New Chance	\$17,820	\$194,640	0.51
Chicago CPC	\$14,090	\$134,150	0.75
Perry Preschool	\$31,840	\$165,430	0.60
Talent Search	\$3,400	\$30,660	3.26

Programs by grade level	Reading ability of target students	Program duration (weeks)	Total cost per student	Literacy domain	Effect size gain	Cost per unit increase in effect size*
Kindergarten average readers:						
K-PALS**	All	20	\$27	Alphabetics	0.61	\$38
Kindergarten struggling readers:						
Stepping Stones	Struggling; behavioral disorders	5	\$479	Alphabetics	0.84	\$570
Sound Partners	20–30th percentile	18	\$791	Alphabetics	0.34	\$2,093
				Fluency	0.48	\$165
First grade struggling readers:						
Fast ForWord Reading 1	Slightly below average	6	\$282	Alphabetics	0.24	\$601
Reading Recovery	Bottom 20th percentile	12–20	\$4,144	Alphabetics	0.70	\$1,480
				Fluency	1.71	\$606
Third grade struggling readers:						
Corrective Reading	Bottom 25th percentile	28	\$10,108	Alphabetics	0.22	\$38,135
				Fluency	0.27	\$6,364
Wilson Reading System	Bottom 25th percentile	28	\$6,696	Alphabetics	0.33	\$13,392

Summary

Main 'problems' with CEA are:

- A. Validity of intervention (lack of specificity/fidelity)
- B. Validity of effects (vagueness)
- C. Helps make decisions

Problems A and B have little to do with CEA *per se*; CEA makes them explicit

Solutions:

- Do more CEA – health research offers only some promise
- Conduct research with decision-makers
- Perform CBA instead of CEA – only eliminate validity of effects problem; change relevance for decision-makers; harder to make comparisons