Figure 1.1 Plot of Hypothetical Data from Three Studies with Grade Level as a Covariate

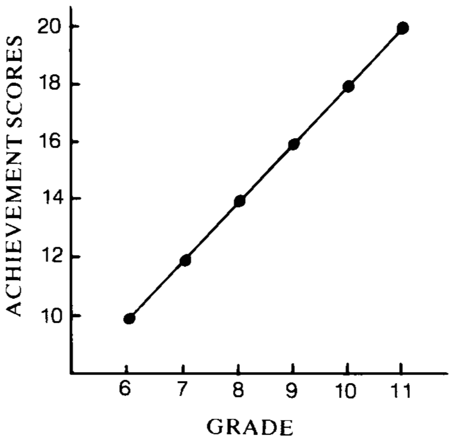

Figure 1.2. Plot of Alternative Hypothetical Data from Three Studies with Grade Level as a Covariate

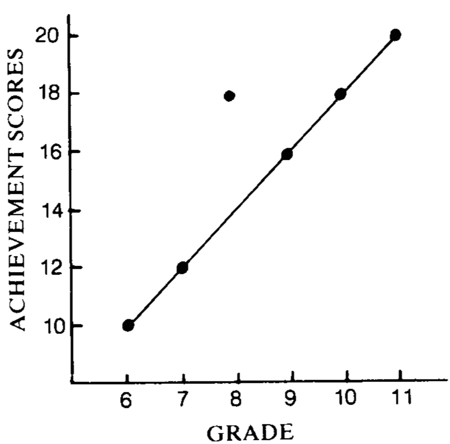

Table 1.1 Measuring Gravitational Force

	Gold	Platinum	Glass
	6.683	6.661	6.678
	6.681	6.661	6.671
	6.676	6.667	6.675
	6.678	6.667	6.672
	6.679	6.664	6.674
	6.672		
Mean	6.678	6.664	6.674
Mean Absolute Deviation	.003	.002	.002
Range	.011	.006	.006

TABLE 3.1 Data Table for Studies with a Pre-treatment and Post-treatment Design

Study Number	Sample Size	Outcome Measure	Pre-treatment Mean	Pre-treatment S.D.	Post-treatment Mean	Post-treatment S.D.	t-Statistic	p-Value	Comments
1	7	Pica-0 %ile	54.30	14.50	52.90	11.00	0.38		My p
6	14	Pica-0 %ile	60.00	19.27	71.00	16.82	5.6224		My $t + p$
11	45	Pica-0	9.08	2.06	11.85	1.70	12.8477		All my calculations
		(Pica-0 %ile)	37.60	18.76	66.51	19.94	13.6985		(Same as above) (Severity collapsed)
14	15	Pica-0	9.42	1.78	11.24	1.79	3.6567		All my calculations
32	104	Pica-0	8.90	s.e.= .22	9.90	s.e.= .22			All values from figures No individual data
37	17	Pica-0 %ile	56.65	19.45	73.18	22.70	4.8866		All my calculations
38	24 moderate	Pica-0	9.91	1.07	10.92	1.22		$<.001$	No individual data
	12 severe	Pica-0	7.87	1.38	8.90	1.20		$<.01$	(Same as above)

TABLE 3.1 *(continued)*

Study Number	Sample Size	Outcome Measure	Pre-treatment Mean	Pre-treatment S.D.	Post-treatment Mean	Post-treatment S.D.	*t*-Statistic	*p*-Value	Comments
54+10	18 individuals	Pica-0 %ile	Difference Score		$\bar{x}$=20	s.d.= 2.30		<.05	Conservative approximation
	16 group	Pica-0 %ile	"	"	$\bar{x}$=15.75	s.d.= 6.02		<.05	Used 24-week completers Stats are anova's
55	29	Pica-0	11.63	1.66	12.23	1.40			on 3 groups
	29	(Pica-0 %ile)	64.07	19.26	70.93	17.03			Time × group × time
62	13	Pica-0	9.48	1.98	12.72	1.50	11.66	<.01	My s.d. calculation only
SS3	2	Pica-0 %ile	45.00	0.63	46.00	1.41	0.07	.25	All my calculations
SS4	2	WAB-AQ	66.00	7.07	70.15	5.16	3.074	.1	All my calculations
SS11	2	Pica-0 %ile	36.00	26.87	43.00	16.97	1.00	.25	All my calculations

Notes: Table is reproduced from Dr. Fromm's notes; s.e. = standard error; s.d. = standard deviation.

Table 3.2 Summary Statistics and Estimates of Effect Size for Studies Using a Pre-treatment and Post-treatment Design

Study Number	Sample Size	Pre-treatment Mean	Pre-treatment S.D.	Post-treatment Mean	Post-treatment r	ES1	ES2
1	7	54.30	14.50	52.90	0.74	– 0.10	– 0.14
6	14	60.00	19.27	71.10	0.92	0.58	1.50
11	45	37.60	18.76	66.51	0.70	1.54	1.92
14	15	9.42	1.78	11.24	0.42	1.02	0.94
32	104	8.90	2.24	9.90	0.74*	0.45	0.62*
37	17	56.65	19.45	73.18	0.79	0.85	1.19
38	36	9.23	1.51	10.25	0.74*	0.68	1.00*
54	34	0.00	18.30*	18.00	0.74*	0.98*	4.04
55	29	64.07	19.26	70.93	0.74*	0.36	0.52*
62	13	9.48	1.98	12.72	0.87	1.64	3.23
SS3	2	45.00	0.63	46.00	0.74*	1.59	0.05
SS4	2	66.00	7.07	70.15	0.74*	0.59	2.18
SS11	2	36.00	26.87	43.00	0.74*	0.26	0.71

TABLE 3.3 Assessment of the Influence of Each Study on the Average Effect Size

Study Number (i)	$\overline{ES1}$	$\overline{wES1}$
1	0.878	0.794
6	0.848	0.696
11	0.742	0.649
14	0.785	0.692
32	0.833	0.930
37	0.799	0.770
38	0.813	0.786
54/10	0.788	0.750
55	0.840	0.815
62	0.733	0.737
SS3	0.738	0.769
SS4	0.821	0.777
SS11	0.837	0.737
median	0.813	0.770

Table 4.1 Studies Reporting PICA Percentiles

Study Number	Sample Size	Pre	Post	%
1	7	54.30	52.90	–1.40
6	14	60.00	71.10	11.10
11	45	37.60	66.51	28.91
37	17	56.65	73.18	16.53
54	34	—	—	18.00
55	29	64.07	70.93	6.86
SS3	2	45.00	46.00	1.00
SS11	2	36.00	43.00	7.00
Average		50.50	60.50	11.02
(weighted)		50.80	67.30	17.03

Table 8.1 Rankings of Research Design Factors' Impact on the Utility of a Study's Results

	Panelists						
	A	B	C	D	E	F	$\bar{x}$
Experimental Manipulation (Definition of Desegregation)	1	1	3	2	1	2	1.6
Experimental Comparison (Adequacy of Control Group)	3	2	2	1	2	1	1.8
Outcome Measure (Measurement of Achievement)	2	6	1	3	4	6	3.6
Population Generality	5	5	4	4	5	5	4.6
Ecological Generality	5	3	5	6	3	4	4.3
Statistical Analysis	4	4	6	5	6	2	4.5

Note : Rank of 1 means factor was judged to have the most impact. The average correlation between rankings was $\bar{r} = +.47$, ranging from $r = -.29$ to $+.77$

Table 8.2 Graduate Reader Beliefs About Desegregation Before and After Reading the Reviews

Question	Before	After
What is the size of the desegregation effect?	4.41 (2.57)	2.58 (1.38)
How confident are you that your belief is accurate?	6.00 (1.92)	7.00 (1.36)
What variables or conditions mediate the effect? (number mentioned)	6.34 (3.45)	8.89 (5.70)

Note: Standard deviations are in parentheses.

Table 8.3 Coding Frame for Moderators of Desegregation Effects

Category		
1.	Attitudes (support)	Of parents, white, black; of students, white, black; of teachers; of school administration; of politicians; of business; of clergy
2.	Student Background	Parent education; family relations; family size; social class, blacks, whites; achievement/ability, blacks, whites; age; self concept/personality
3.	Community Background	Area of country; urban vs. rural; size; historical conditions; media presentation
4.	School Characteristics	Size; location; quality (curriculum)
5.	Classroom Characteristics	Size (teacher/pupil ratio; number of teachers); seating (interaction) patterns; open vs. traditional
6.	Teacher Characteristics	Ability; labeling (expectation) effects; training for desegregation; race
7.	Achievement Definition	Measurement; subject matter; interpretation of measurement
8.	Desegregation Definition	Voluntary vs. forced; community involvement in implementation; need (distance) for busing; black/white ratio; length of implementation; who gets bused
9.	Money (expenditures)	Resources for teachers; other support services (personnel)
10.	Change of School Problems (confusion)	Preparation of students
11.	Other	

Table 8.4 Change in the Number of Moderators Mentioned by Readers in Eleven Categories

Category	Prior	Post	t-value	p-value
1. Attitudes	1.71	1.86		
2. Student Background	1.11	1.39		
3. Community Background	0.86	0.54		
4. School Characteristics	0.32	0.38		
5. Classroom Characteristics	0.11	0.25		
6. Teacher Characteristics	0.46	0.86		
7. Achievement Definition	0.14	0.79	3.23	.007
8. Desegregation Definition	0.96	1.75	2.27	.040
9. Money	0.46	0.25		
10. Change of School Problems	0.07	0.25	2.11	.060
11. Other	0.14	0.57		

Table 8.5 Relations Between Reader's Background and Prior and Post Beliefs About Desegregation

	Topic Familiarity	Research Expertise	Political Beliefs	Prior Effect Size	Prior Confidence	Prior Moderators	Post Effect Size	Post Confidence	Post Moderators
Topic Familiarity	—	−.04	−.02	−0.47*	.48*	−.30	.35	.38*	−.24
Research Expertise		—	.05	−.01	−.06	.10	−.45*	.00	.54*
Political Beliefs			—	.01	.08	−.36	.39	−.35	.15
Prior Effect Size				—	.12	.51*	.01	−.53*	.07
Prior Confidence					—	.06	.48*	.09	−.38
Prior Moderators						—	−.12	−.19	.58*
Post Effect Size							—	−.22	−.38
Post Confidence								—	−.26
Post Moderators									—

Notes: All tests are based on between 9 and 13 degrees of freedom. More liberal beliefs were given higher numerical values. Effect sizes were coded as negative values if the reader thought desegregation had negative effects.

* $.15 > p > .05$

Table 9.1 Integrative Review Conceptualized as a Research Project

	Stage of Research	
Stage Characteristics	Problem Formulation	Data Collection
Research Question Asked	What evidence should be included in the review?	What procedures should be used to find relevant evidence?
Primary Function in Review	Constructing definitions that distinguish relevant from irrelevant studies.	Determining which sources of potentially relevant studies to examine.
Procedural Differences That Create Variation in Review Conclusions	1. Differences in included operational definitions. 2. Differences in operational detail.	Differences in the research contained in sources of information.
Sources of Potential Invalidity in Review Conclusions	1. Narrow concepts might make review conclusions less definitive and robust. 2. Superficial operational detail might obscure interacting variables.	1. Accessed studies might be qualitatively different from the target population of studies. 2. People sampled in accessible studies might be different from target population of people.

Source: Cooper (1982).

Stage of Research		
Data Evaluation	Analysis and Interpretation	Public Presentation
What retrieved evidence should be included in the review?	What procedures should be used to make inferences about the literature as a whole?	What information should be included in the review report?
Applying criteria to separate "valid" from "invalid" studies.	Synthesizing valid retrieved studies.	Applying editorial criteria to separate important from unimportant information.
1. Differences in quality criteria. 2. Differences in the influence of nonquality criteria.	Differences in rules of inference.	Differences in guidelines for editorial judgment.
1. Nonquality factors might cause improper weighing of study information. 2. Omissions in study reports might make conclusions unreliable.	1. Rules for distinguishing patterns from noise might be inappropriate. 2. Review-based evidence might be used to infer causality.	1. Omission of review procedures might make conclusions irreproducible. 2. Omission of review findings and study procedures might make conclusions obsolete

Table 9.2 Authors' Ranking of Various Categories to Describe Their Reviews

		Percentage of Authors Using as Descriptor		
Characteristic	Categories	Primary	Secondary	Omitted
Focus	Research outcomes	56%	12%	25%
	Research methods	9	17	47
	Theories	19	23	32
	Practices or applications	22	23	26
Goal	Integration			
	(a) Generalization	23	19	35
	(b) Conflict resolution	7	11	59
	(c) Linguistic bridge-building	8	6	68
	Criticism	42	16	24
	Identification of central issues	24	21	34
Perspective	Neutral representation	81	4	14
	Espousal of position	18	18	65
Coverage	Exhaustive	37	2	52
	Exhaustive with representative citation	21	4	70
	Exhaustive with central citation	14	7	75
	Representative	19	6	70
	Central or pivotal	7	6	78
Organization	Historical	7	9	77
	Conceptual	76	6	17
	Methodological	15	13	70
Audience	Specialized scholars	39	26	28
	General scholars	34	23	27
	Practitioners	31	18	28
	Policymakers	4	5	72
	General public	2	5	77

Table 10.1 Four Steps in the Evaluation Synthesis Methodology

Major Steps	Subtasks
A. Formulate and Define Questions	1. Identify issues 2. Review literature and legislative history 3. Conduct interviews a. researchers b. professionals and public interest groups c. federal agency staff
B. Identify and Collect Relevant Studies	1. Conduct bibliographic searches 2. Conduct interviews a. researchers b. professionals and public interest groups c. federal agency staff
C. Assess Usefulness and Reliability of Studies	1. Classify by study questions 2. Review studies a. research design b. methodology (1) sampling (2) data collection (3) statistical analysis c. findings 3. Summarize
D. Synthesize Information	1. Assess studies and findings a. strengths and weaknesses b. similarities and trends c. generalizability 2. Relate findings to synthesis questions 3. Summarize what is known 4. Have experts review

Table 11.1 Binomial Effect Size Display (BESD) for an *r* of .32 that Accounts for "Only" 10 Percent of the Variance

	Treatment Result		
Condition	Alive	Dead	Σ
Treatment	66%	34%	100%
Control	34	66	100
Σ	100	100	200

Table 11.2 Changes in Success Rates (BESD) Corresponding to Various Values of r^2 and r

Effect Sizes		Equivalent to a Success Rate Increase		
r^2	r	From	To	Difference in Success Rates[a]
.00	.02	.49	.51	.02
.00	.04	.48	.52	.04
.00	.06	.47	.53	.06
.01	.08	.46	.54	.08
.01	.10	.45	.55	.10
.01	.12	.44	.56	.12
.03	.16	.42	.58	.16
.04	.20	.40	.60	.20
.06	.24	.38	.62	.24
.09	.30	.35	.65	.30
.16	.40	.30	.70	.40
.25	.50	.25	.75	.50
.36	.60	.20	.80	.60
.49	.70	.15	.85	.70
.64	.80	.10	.90	.80
.81	.90	.05	.95	.90
1.00	1.00	.00	1.00	1.00

[a]The difference in success rates in a BESD is identical to r.

Table 11.3 Binomial Effect Size Display for the Discontinued Propranolol Study

	Treatment Result		
Condition	Alive	Dead	Σ
Propranolol	52%	48%	100%
Placebo	48	52	100
Σ	100	100	200