Graphical Methods for Meta-analysis: Looking at the forest and the trees

Hannah Rothstein
Baruch College
& Biostat, Inc
SIOP, April 2003

Acknowledgements

I would like to thank Brian McNatt and Tim Judge for supplying the unpublished information necessary for me to conduct these analyses

Use of their data in this presentation is in no way meant to be critical of their meta-analyses.

Key Points

- Graphical devices are important and underutilized tools for presenting meta-analytic results
- Properly used, they simplify, clarify, and amplify text-based and tabular presentations
- Forest plots
- Funnel plots

Key Points

- The data I am showing for today are uncorrected for artifacts
- Forest and funnel plots can and should be constructed for corrected coefficients, but only if corrections have been made to the **individual** effects
- They cannot be used with the artifact distribution method

Original Information
(from Table 3, McNatt 2000)

Overall Effect: Effects of Pygmalion interventions on Management Performance

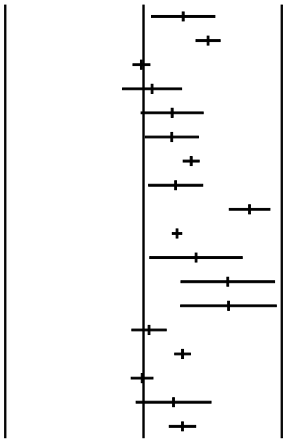
N. Studies	N	K	D uncor	D corr	Est.pop.SD	% var accted for	80% cred value
17	2874	58	.99	1.13	.77	6	.14-2.22

McNatt's Original Information -2

- What was the pattern of results?
- How much did the primary studies agree or disagree?
  - Were there any outliers?
  - How many effects were positive?
  - How many negative?
  - How much overlap was there in the confidence intervals?

Forest Plot of McNatt (2000) Pygmalion Effects, Sorted By Author

Citation	Effect	StdErr	NTotal	PValue
Crawford et al.	1.150	.446	28	.011
Davidson & Eden	1.870	.177	225	.000
Dvir et al Sample 1	-.060	.125	315	.631
Dvir et al Sample 2	.250	.414	26	.550
Dvir et al Sample 3	.830	.436	27	.060
Dvir et al Sample 4	.820	.377	35	.031
Eden & Davidson	1.380	.117	360	.000
Eden & Ravid	.930	.383	31	.016
Eden & Shani	3.070	.295	105	.000
Eden 1990b	.970	.070	1000	.000
King Sample 1	1.520	.628	17	.018
King Sample 2	2.440	.645	20	.000
King Sample 3	2.460	.656	19	.000
Mase	.160	.251	87	.525
Oz & Eden	1.130	.115	350	.000
Sutton & Woodman	-.040	.160	182	.803
Vrugt	.870	.508	17	.094
Combined (17)	**1.129**	**.195**	**2844**	**.000**



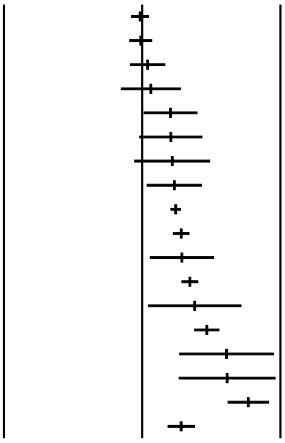
Data are standardized mean differences (ds) uncorrected for artifacts

Forest Plot of McNatt (sorted by author)

- What new information is contained in this graph?
  - Almost all the effects are positive
  - There is a substantial amount of variability among the effects
  - Note: The amount of variability could change some if artifact corrections are made

Forest Plot of McNatt (2000) Pygmalion Effects, Sorted By Effect

Citation	Effect	StdErr	NTotal	PValue
Dvir et al Sample 1	-.060	.125	315	.631
Sutton & Woodman	-.040	.160	182	.803
Mase	.160	.251	87	.525
Dvir et al Sample 2	.250	.414	26	.550
Dvir et al Sample 4	.820	.377	35	.031
Dvir et al Sample 3	.830	.436	27	.060
Vrugt	.870	.508	17	.094
Eden & Ravid	.930	.383	31	.016
Eden 1990b	.970	.070	1000	.000
Oz & Eden	1.130	.115	350	.000
Crawford et al.	1.150	.446	28	.011
Eden & Davidson	1.380	.117	360	.000
King Sample 1	1.520	.628	17	.018
Davidson & Eden	1.870	.177	225	.000
King Sample 2	2.440	.645	20	.000
King Sample 3	2.460	.656	19	.000
Eden & Shani	3.070	.295	105	.000
Combined (17)	**1.129**	**.195**	**2844**	**.000**



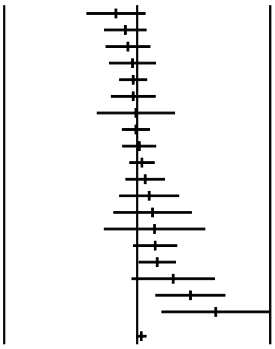
Data are standardized mean differences (ds) uncorrected for artifacts

Forest Plot of McNatt (by effect size)

- What new information is contained in this graph?
- How quickly can you grasp the information being relayed?

Forest Plot of Raudenbush (1984) Expectancy Effects, Sorted By Effect Size

Citation	NTotal	Effect	StdErr	Lower	Upper
Claiborn 1969	125	-.320	.221	-.758	.118
Fine 1972	159	-.180	.159	-.494	.134
Jose & Cody 1971	144	-.140	.167	-.470	.190
Ginsburg 1970	132	-.070	.174	-.415	.275
Evans & Rosenthal 1968	477	-.060	.103	-.263	.143
Greiger 1970	144	-.060	.167	-.390	.270
Keschock 1970	48	-.020	.289	-.601	.561
Fielder et al 1971	746	-.020	.103	-.223	.183
Rosenthal 1974	416	.030	.126	-.218	.278
Fleming & Anttonen 1971	457	.070	.094	-.114	.254
Conn 1968	258	.120	.147	-.170	.410
Flowers 1966	81	.180	.223	-.264	.624
Henrikson 1970	51	.230	.291	-.354	.814
Pellengrini 1972 2	33	.260	.371	-.496	1.016
Kester 1969	149	.270	.165	-.055	.595
Rosenthal & Jacobson 1968	320	.300	.139	.026	.574
Carter 1970	44	.540	.307	-.079	1.159
Mazwell 1970	64	.800	.260	.281	1.319
Pellengrini 1972 1	33	1.180	.397	.371	1.989
Combined (19)	**3881**	**.061**	**.036**	**-.011**	**.132**



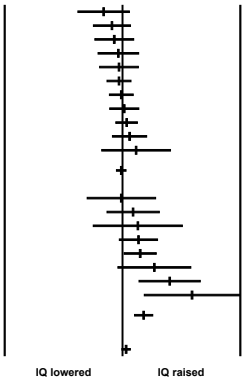
Data are standardized mean differences (ds) uncorrected for artifacts

Grouping Data

- Should we conclude there is no effect?
  - The overall effect is small
  - The standard error is not huge
  - But, there seems to be a pattern here

Raudenbush (1984) Expectancy Data Grouped By Contact, Sorted by ES

contact	Citation	NTotal	Effect	StdErr	Lower	Upper
more than one week	Claiborn 1969	125	-.320	.221	-.758	.118
more than one week	Fine 1972	159	-.180	.159	-.494	.134
more than one week	Jose & Cody 1971	144	-.140	.167	-.470	.190
more than one week	Ginsburg 1970	132	-.070	.174	-.415	.275
more than one week	Greiger 1970	144	-.060	.167	-.390	.270
more than one week	Evans & Rosenthal 1968	477	-.060	.103	-.263	.143
more than one week	Fielder et al 1971	746	-.020	.103	-.223	.183
more than one week	Rosenthal 1974	416	.030	.126	-.218	.278
more than one week	Fleming & Anttonen 1971	457	.070	.094	-.114	.254
more than one week	Conn 1968	258	.120	.147	-.170	.410
more than one week	Henrikson 1970	51	.230	.291	-.354	.814
more than one week (11)		**3109**	**-.021**	**.041**	**-.102**	**.059**
one week or less	Keschock 1970	48	-.020	.289	-.601	.561
one week or less	Flowers 1966	81	.180	.223	-.264	.624
one week or less	Pellengrini 1972 2	33	.260	.371	-.496	1.016
one week or less	Kester 1969	149	.270	.165	-.055	.595
one week or less	Rosenthal & Jacobson 1968	320	.300	.139	.026	.574
one week or less	Carter 1970	44	.540	.307	-.079	1.159
one week or less	Mazwell 1970	64	.800	.260	.281	1.319
one week or less	Pellengrini 1972 1	33	1.180	.397	.371	1.989
one week or less (8)		**772**	**.358**	**.078**	**.204**	**.512**
Combined (19)		**3881**	**.061**	**.036**	**-.011**	**.132**



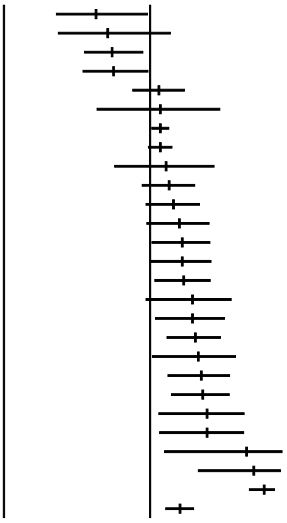
Data are standardized mean differences , uncorrected for artifacts

Raudenbush Moderators

- When sorted by length of contact with students, a clear moderator effect appears
  - The high prior contact group shows no effects
  - The low prior contact group shows an effect of about .36

Irvin t al. (1999, JCCP) Relapse Prevention Meta-Analysis

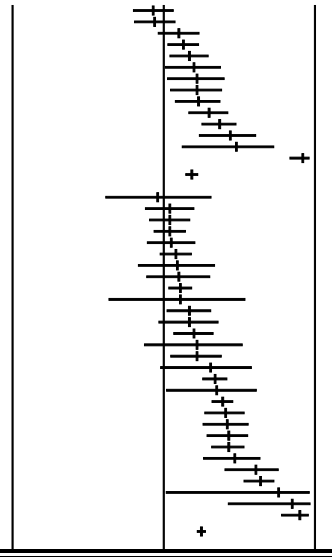
Substance	Citation	Effect	StdErr	Lower	Upper	NTotal
Tobacco	Supnick & Coletti	-.370	.189	-.640	-.018	31
Cocaine	Carroll	-.290	.224	-.627	.139	23
Cocaine	Wells et al	-.260	.110	-.448	-.050	85
Tobacco	Sobell et al	-.250	.123	-.459	-.014	69
Tobacco	Hall et al	.060	.091	-.118	.235	123
Alcohol	O-Connell	.070	.229	-.362	.478	22
Tobacco	Stevens et al	.070	.030	.011	.128	1119
Tobacco	Stevens & Hollis	.070	.041	-.011	.150	587
Polysubstance	Ashkanazi	.110	.183	-.242	.437	33
Tobacco	Zelman et al.	.130	.094	-.054	.305	116
Polysubstance	Hawkins et al 1989	.160	.097	-.028	.337	110
Tobacco	Gregory	.200	.115	-.022	.403	79
Tobacco	Goldstein et al	.220	.108	.012	.409	89
Tobacco	Hill et al	.220	.113	.003	.417	82
Cocaine	Carroll et al.	.230	.103	.032	.411	97
Alcohol	Chaney et al.	.290	.167	-.028	.555	39
Alcohol	O-Farrell et al.	.290	.134	.037	.508	59
Polysubstance	Knight et al.	.310	.104	.116	.481	95
Alcohol	Annis & Peachey	.330	.167	.016	.585	39
Alcohol	O-Malley et al	.350	.124	.122	.543	68
Polysubstance	Hawkins et al 1986	.360	.117	.146	.542	76
Alcohol	Maisto et al	.390	.180	.060	.643	34
Alcohol	Sandahl & Ronnberg	.390	.177	.065	.640	35
Alcohol	Ito et al	.660	.354	.100	.903	11
Polysubstance	Peters et al.	.710	.277	.331	.892	16
Alcohol	Kranzler et al	.780	.110	.680	.852	85
Combined (26)		**.204**	**.051**	**.107**	**.298**	**3222**



Grouping and Sorting

T. Judge Height-Perf. Grouped by Gender Sorted by ES

gend	Citation	NTotal	Effect	Lower	Upper
female	Reynolds 15b	219	-.07	-.20	.06
female	Villimez 29 female	213	-.06	-.19	.08
female	Moore 14	207	.10	-.04	.23
female	Smoll 27 female	357	.13	.03	.23
female	Melamed 13 female	227	.17	.04	.29
female	Eisenberg 32 ii	108	.20	.01	.37
female	Eisenberg 32 i	100	.22	.02	.40
female	Roth 30	123	.22	.04	.38
female	Werner 19 female	159	.23	.08	.37
female	Lerner 11b	192	.30	.17	.42
female	Lichtman 26	227	.37	.25	.48
female	Bellingrath 2b female	74	.44	.24	.61
female	Eisenberg 32 iii	27	.48	.12	.73
female	Lechelt 10 female	29	.92	.84	.96
female (14)		**2262**	**.19**	**.15**	**.23**
male	Phillips 33	32	-.04	-.38	.31
male	Evans 3	149	.04	-.12	.20
male	Reynolds 15a	218	.04	-.09	.17
male	Smoll 27 male	357	.04	-.06	.14
male	Sheldon 23	155	.05	-.11	.21
male	Mazur 12	356	.08	-.02	.18
male	Kitson 8	60	.09	-.17	.34
male	Bellingrath 2a male	87	.10	-.11	.30
male	Hoobler 7	656	.11	.03	.18
male	Zeleny 22 i	19	.11	-.36	.54
male	Villimez 29 male	175	.17	.02	.31
male	Azen 1	95	.17	-.03	.36
male	Melamed 13 male	208	.20	.07	.33
male	Min 31	34	.22	-.13	.52
male	Willoughby 34	125	.22	.05	.38
male	Eisenberg 32 male	36	.31	-.02	.58
male	Bonuso 25	472	.34	.26	.42
male	Zeleny 22 ii	35	.35	.02	.61
male	Talbert	578	.39	.32	.46
male	Werner 19 male	159	.41	.27	.53
male	Lerner 11a	119	.42	.26	.56
male	Partridge 24 i	143	.43	.29	.55
male	Partridge 24 ii	226	.43	.32	.53
male	Prieto 35	69	.47	.26	.64
male	Hensley 6	52	.61	.40	.76
male	Kurtz 9	140	.64	.53	.73
male	Collins 28	7	.76	.02	.96
male	Young 21	9	.85	.43	.97
male	Lechelt 10 male	24	.90	.78	.96
male (29)		**4795**	**.25**	**.22**	**.28**



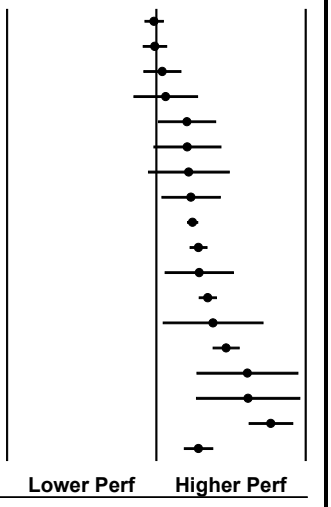
Multiple Page Graphs

Constructing a Forest Plot-1

- It is a graphical display
- It depicts the individual study effects + Cis
- Study symbols
  - Constant size symbols for each study
  - Proportional symbols for each study.
    - The size of the symbol shows the % weight each study has in the pooled analysis

McNatt, Constant Size Symbols

Citation	Effect	StdErr	NTotal	PValue
Dvir et al Sample 1	-.060	.125	315	.631
Sutton & Woodman	-.040	.160	182	.803
Mase	.160	.251	87	.525
Dvir et al Sample 2	.250	.414	26	.550
Dvir et al Sample 4	.820	.377	35	.031
Dvir et al Sample 3	.830	.436	27	.060
Vrugt	.870	.508	17	.094
Eden & Ravid	.930	.383	31	.016
Eden 1990b	.970	.070	1000	.000
Oz & Eden	1.130	.115	350	.000
Crawford et al.	1.150	.446	28	.011
Eden & Davidson	1.380	.117	360	.000
King Sample 1	1.520	.628	17	.018
Davidson & Eden	1.870	.177	225	.000
King Sample 2	2.440	.645	20	.000
King Sample 3	2.460	.656	19	.000
Eden & Shani	3.070	.295	105	.000
Combined (17)	**1.129**	**.195**	**2844**	**.000**



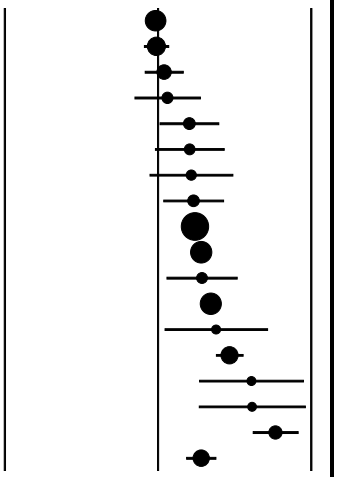
Data are standardized mean effects , uncorrected for artifacts

Perceptually Misleading?

- The eye is drawn to the longer error bars (CI bars)
- Less informative studies have a relatively greater visual effect
- The use of proportional symbols for point estimates reduces this distortion

McNatt, Proportional Symbols (Telescoped)

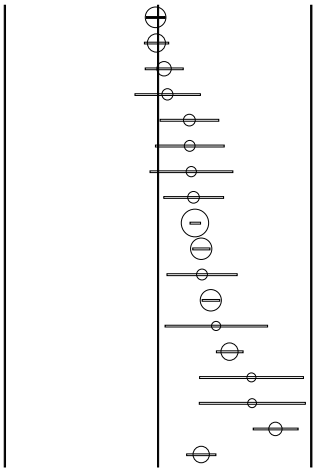
Citation	Effect	StdErr	NTotal	PValue
Dvir et al Sample 1	-.060	.125	315	.631
Sutton & Woodman	-.040	.160	182	.803
Mase	.160	.251	87	.525
Dvir et al Sample 2	.250	.414	26	.550
Dvir et al Sample 4	.820	.377	35	.031
Dvir et al Sample 3	.830	.436	27	.060
Vrugt	.870	.508	17	.094
Eden & Ravid	.930	.383	31	.016
Eden 1990b	.970	.070	1000	.000
Oz & Eden	1.130	.115	350	.000
Crawford et al.	1.150	.446	28	.011
Eden & Davidson	1.380	.117	360	.000
King Sample 1	1.520	.628	17	.018
Davidson & Eden	1.870	.177	225	.000
King Sample 2	2.440	.645	20	.000
King Sample 3	2.460	.656	19	.000
Eden & Shani	3.070	.295	105	.000
Combined (17)	**1.129**	**.195**	**2844**	**.000**



Data are standardized mean effects , uncorrected for artifacts

McNatt, Proportional Symbols (Telescoped)

Citation	Effect	StdErr	NTotal	PValue
Dvir et al Sample 1	-.060	.125	315	.631
Sutton & Woodman	-.040	.160	182	.803
Mase	.160	.251	87	.525
Dvir et al Sample 2	.250	.414	26	.550
Dvir et al Sample 4	.820	.377	35	.031
Dvir et al Sample 3	.830	.436	27	.060
Vrugt	.870	.508	17	.094
Eden & Ravid	.930	.383	31	.016
Eden 1990b	.970	.070	1000	.000
Oz & Eden	1.130	.115	350	.000
Crawford et al.	1.150	.446	28	.011
Eden & Davidson	1.380	.117	360	.000
King Sample 1	1.520	.628	17	.018
Davidson & Eden	1.870	.177	225	.000
King Sample 2	2.440	.645	20	.000
King Sample 3	2.460	.656	19	.000
Eden & Shani	3.070	.295	105	.000
Combined (17)	**1.129**	**.195**	**2844**	**.000**



Data are standardized mean effects , uncorrected for artifacts

Constructing a Forest Plot-3

- A vertical line drawn down the middle of the graph represents no effect
  - For standardized mean effects and correlations, this is represented by a zero
  - For odds ratios, this is represented by a one
- The horizontal line at the bottom is the scale measuring the magnitude of the effect

Funnel Plots

- A means for detecting publication biasGraphs effect size by effect precision
  - Magnitude of the effect
  - Sample size or 1/standard error

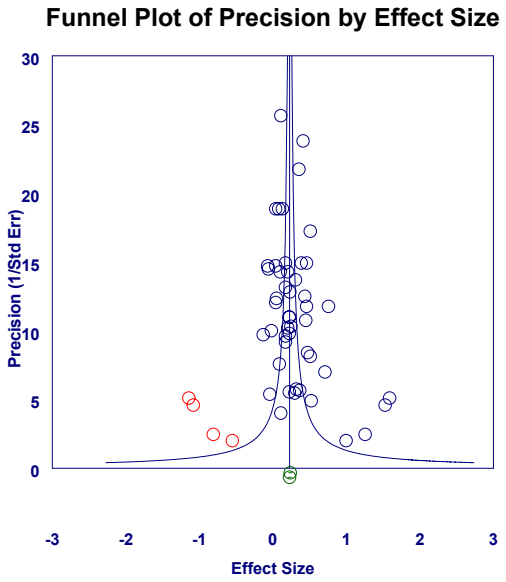
Funnel Plots 2

- If the meta analysis had captured all the relevant studies we would expect the funnel plot to be symmetric.
- If the funnel plot is actually asymmetric,
  - Relatively high number of small studies falling toward the right (representing a large treatment effect)
  - Relatively few falling toward the left,
  - We are concerned that these left-hand studies exist, but are missing from the analysis.

T. Judge Height-Perf Data

Effects shown are Fisher's zs.

Red circles denote imputed studies



3/31/2003

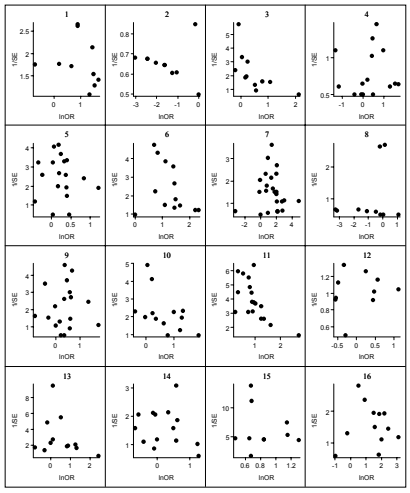
Funnel Plots 4

- Trim and Fill Method
- Suggests that there are 4 "missing" studies
- However, the change in the mean correlation is only about .01
- We can thus feel confident in our results
- Caveat: These are uncorrected effects, the shape of the funnel can change when you make corrections for artifacts.

Funnel Plots 5

- The Judge Height-Performance funnel plot is nicely behaved
- This is partly because it is based on 48 ES
- In Funnel Plots based on fewer data points (more common in Medicine, and in moderator analyses), symmetry and asymmetry are harder to detect, as shown in the next slide

Data courtesy of Alex Sutton, from Sutton et al. (2000) BMJ



Funnel Plots 7

- Funnel plot analysis should probably not be used when there are only a few studies in the meta-analysis
- This is for the same reason moderator analyses should not be done when there is a small K
- Second order sampling error

Funnel Plots 8

- While small studies may show relatively large effects because of publication bias, they may also show relatively large effects for other reasons
- Perhaps researchers are better able to control the conditions of small studies than large studies

Funnel Plots 9

- However, whatever the basis for the bias, the essential goal of this analysis, to determine whether or not we can have confidence in the results, remains valid.

In Conclusion

- Demonstration of the added value of graphic methods
- Encouragement of their wider use in I/O meta-analyses