

## Index

---

- Abelson's paradox 43(note 7)  
 accidental findings 78, 135  
 AERA Standards for Reporting 5, 19, 137  
 alpha ( $\alpha$ ) 48, 49–52, 54, 55, 69(note 13, note 15), 82  
   adjusting 79, 82, 85(note 16), 135, 136  
   arguments against adjusting 84(note 10)  
   statistical power and 50, 56, 57  
 alpha-to-beta ratio, *see* beta-to-alpha ratio  
 alternative hypothesis 47  
 alternative plausible explanations 21, 39–40  
 Alzheimer's study 4, 5, 9, 40, 47, 57–58, 59, 60, 70(note 16), 82, 110, 111  
 APA *Publication Manual* 4, 5, 19, 25(note 2), 137  
 Apophis asteroid 36  
 a priori power analysis, *see* power analysis;  
   prospective  
 arbitrary scales 32–35  
 Asian financial crisis 35, 43(note 4)  
 aspirin study 23, 24, 52  
 astrological study 51  
 astronomer, the foolish 47  
 availability bias xiv, 117, 119, 121, 122, 125, 133(note 10), 134  
   how to detect 120  
  
 Beijing Olympics 36  
 beta ( $\beta$ ) 49–52, 55, 61  
 beta-to-alpha ratio 50, 53, 54–55, 56, 69(note 11, note 13), 79–80, 82, 84(note 12), 136  
 binomial effect size display 21, 23–24, 30(note 24)  
 bogus manuscript study 119  
 Bonferroni correction 79, 84(note 9)  
  
 Challenger explosion 56  
 Coding 98–101  
   the drudgery of 101  
   interrater agreement 101, 114(note 11)  
 coefficient of determination ( $r^2$ ) 12, 13  
 coefficient of multiple determination ( $R^2$ ) 12, 13, 15, 27(note 10, note 11)  
 coefficient of multiple determination, adjusted ( $_{adj}R^2$ ) 12, 13, 15  
  
 Cohen's  $d$  10, 12, 13, 15, 21, 40  
 Cohen's effect size benchmarks 33, 40–42, 93  
   criticisms of 41–42  
 Cohen's  $f$  12, 13, 15, 31  
 Cohen's  $f^2$  12, 13  
 Cohen's power recommendations 53–54  
 common language effect size 21–22  
 confidence intervals 17–21, 65, 66, 70(note 18), 92, 136  
   central vs. non-central 19, 21  
   credibility intervals vs. 106  
   defined 17, 18  
   editorial calls for 19, 29(note 18)  
   graphing 21  
   hypothesis testing and 18, 104, 107  
   methods for constructing 19–21  
   misuse of 17  
 CONSORT statement 39, 72(note 24)  
 correlation coefficient  $r$  11, 16, 22  
   *see also* part correlation, partial correlation, Pearson product moment correlation, phi coefficient, point-biserial correlation, semipartial correlation, Spearman rank correlation  
 correlation matrix 16, 59, 100, 136  
 correlation ratio, *see* eta squared ( $\eta^2$ )  
 Cramér's contingency coefficient  $V$  11, 13, 15  
 credibility interval 107  
 Cuban missile crisis 36, 40  
 Cydonian Face 51  
  
*d*, *see* Cohen's  $d$   
 databases, bibliographic 97  
 differences between groups, *see* effect size,  $d$ -family  
 directional test, *see* one-tailed test  
 dust-mite study 130  
  
 effect xiii, 4, 47, 52, 134  
   *see also* small effects  
 effect size 4–6, 48, 65, 93, 95, 121, 134  
   calculators 14, 28(note 14)  
   corrected vs uncorrected estimates 27–28(note 12)  
    $d$ -family 7–11, 13, 16, 99  
   estimation of 5, 6, 12, 43(note 3)  
   index 6, 16, 26(note 3), 136

- minimum detectable 57, 60, 63–64  
 observed vs. population effect size 5, 18, 38, 59, 60, 70(note 17), 73, 104, 106, 127  
*r*-family 11–12, 13, 16, 99  
 SPSS calculations 12, 15, 27(note 11)
- effect size reporting 16, 21–24  
 editorial calls for xiv, 4, 24, 25(note 1), 25–26(note 2)
- epsilon squared 12, 13
- equations for,  
 confidence intervals 17, 105  
 converting odds to probabilities 26(note 5)  
 converting probabilities to odds 26(note 5)  
 fail-safe *N* 133(note 5)  
 margin of error 20  
*Q* statistic 107, 143, 147  
 standard error 20  
 transforming chi-square to *r* 28(note 15)  
 transforming *d* to *r* (equal groups) 16  
 transforming *d* to *r* (unequal groups) 28(note 15)  
 transforming *d* to *z* 133(note 5)  
 transforming *r* to *d* 16  
 transforming *r* to *z* 146  
 transforming *z* to *r* 28(note 15), 148  
 variance 106, 146  
 variance (between studies) 144  
 variance (combined) 144  
 variance (within studies) 142  
 weighted mean effect size – *d* (FE) 142  
 weighted mean effect size – *d* (RE) 145  
 weighted mean effect size – *r* (FE) 147  
 weighted mean effect size – *r* (RE) 148  
 weighted mean effect size – Hunter and Schmidt 103
- eta squared ( $\eta^2$ ) 12, 13, 15
- experimentwise error rate, *see* familywise error rate
- fail-safe *N* 122, 133(note 5)  
 Rosenthal's threshold 122
- false negative 48, 56, 82, 124, 130
- false positive 48, 51, 56, 80, 119, 124, 136
- familywise error rate 78, 79, 124
- file drawer problem 91, 117–119
- fishing 78, 84(note 8)
- five-eighty convention 54, 80
- fixed-effects procedures, *see* meta-analysis; fixed effects procedures
- funnel plot 120–122
- gender and map-reading study 141
- Glass's delta ( $\Delta$ ) 10, 13, 15
- global financial crisis 35
- Goodman and Kruskal's lambda ( $\lambda$ ) 11, 12, 13, 15
- HARKing (hypothesizing after the results are known) 78, 80, 84(note 8), 124
- Hedges' *g* 10, 13, 15, 27(note 9)
- Hong Kong flu 35
- hormone replacement therapy 37
- interpretation xiv, 5, 6, 16, 35–43, 65, 108–109, 137  
 contribution to theory 38–40, 109  
 editorial calls for 39–40, 42, 108  
 in the context of past research 25(note 2), 38–39  
 the problem of 31, 32–35, 48, 90, 91, 94, 109  
 statistical significance and 4, 6, 16, 32, 42, 95  
*see also* thresholds for interpreting effect sizes
- intrater reliability, *see* coding; interrater agreement
- Kendall's tau ( $\tau$ ) 11, 13, 15
- kyptonite meta-analysis 102–104
- literature review, *see* meta-analysis, narrative review
- logit coefficient 12, 15
- logged odds ratio, *see* logit coefficient
- magnesium study 121, 125
- margin of error 18, 20, 134
- market orientation meta-analyses 90, 96–97, 111, 114(note 9)
- measurement error 66, 81, 85(note 14), 95, 135  
 reliability 66, 134
- measures of association, *see* effect size, *r*-family meta-analysis 61, 90, 94–97, 112, 115(note 18), 132, 134  
 advantages of 96–97, 111  
 apples and oranges problem 98, 101, 111, 114(note 9)  
 bias affecting 117, 126, 127  
 collecting studies for 97–98  
 combining effect sizes 18, 100, 125  
 confidence intervals in 93, 105, 106, 127, 129, 151  
 defined 94  
 eligibility criteria 98  
 fixed-effects procedures 128–130, 137, 141  
 garbage in, garbage out criticism 123  
 Hedges et al. method 109, 131, 141  
 history of 95, 96  
 homogeneity of the sample 107–108, 129, 131, 143, 149  
 Hunter and Schmidt method 109, 131, 141, 149–152  
 information overload and 112  
 large scale randomized control trials vs. 116  
 mean effect size 93, 95, 103, 137, 149  
 measurement error and 100, 103, 151  
 mixing good and bad results 124–127  
 mixing good and bad studies 123–124  
 moderator analysis 100, 108, 111  
 procedures for 109, 141–148, 150  
 random-effects procedures 128–130, 137, 144  
 replication research and 109–111  
 statistical power of 123, 125, 126, 127, 130–131  
 theory development and 111–112  
*see also* availability bias, coding, file drawer problem, reviewer bias
- meta-analytic thinking 93
- minimum detectable effects, *see* effect size: minimum detectable

172      **The Essential Guide to Effect Sizes**

- multiplicity curse, *see* multiplicity problem
- multiplicity problem 71(note 24), 78, 79, 124
- narrative review xvi, 90, 91–92  
 limitations of 94, 96
- narrative summary, *see* narrative review
- National IQ Test 15, 31, 94
- nondirectional test, *see* two-tailed test
- nonresponse bias 71(note 24)
- nonsignificant results 58–60, 71(note 19), 92, 100, 110, 119, 120, 136  
 misinterpreting 32, 52, 59
- null hypothesis 47–48, 49, 50, 60, 65, 66, 67(note 1), 68(note 4), 134
- null hypothesis significance testing, *see* statistical significance testing
- odds ratio 7–9, 13, 15, 27(note 10, note 11)
- omega squared 12, 13
- omnibus effect 63, 100
- one-tailed test 70(note 15), 81
- overpowered tests 52, 53
- part correlation 15, 27(note 11), 63
- partial correlation 15, 27(note 11)
- partial eta-squared 15
- Pearson's contingency coefficient  $C$  11, 13, 15
- Pearson product moment correlation  $r$  11, 13, 15
- phi coefficient ( $\phi$ ) 11, 13, 15, 27(note 10)
- polio vaccine 23
- point-biserial correlation ( $r_{pb}$ ) 11, 13, 15, 16
- post hoc hypotheses 79, 135, 137
- post hoc power analysis, *see* power analysis, retrospective
- POV, *see* proportion of shared variance
- power, *see* statistical power
- power analysis 47, 56–61, 62, 67, 73, 127  
 prospective 57–58, 60, 65, 110, 131, 134  
 retrospective 58–61, 73  
 SPSS and 60
- power calculators 62, 71(note 22)
- power surveys 73–74, 76, 77  
*see also* statistical power of published research
- practical significance, *see* significance; practical
- precision 5, 8, 17–21, 29(note 19), 64, 66, 92, 93, 120, 134, 136
- Premarin study 37
- preventive medicine 37
- probability of superiority ( $PS$ ) 13, 21, 22
- propranolol study 36, 37
- proportion of shared variance 11, 12, 22
- prospective power analysis, *see* power analysis; prospective
- psychotherapy meta-analyses 95, 96
- publication bias 55, 80, 91, 101, 119–120, 132(note 1)
- $p$  values 16, 48, 49, 68(note 3), 69(note 15), 134, 136  
 and the likelihood of publication 83(note 4)  
 and statistical power 50, 60  
 limitations of 16, 18–19, 29(note 18), 42, 49, 54, 119
- vs. effect sizes 33, 52, 53, 92, 100, 136  
*see also* alpha ( $\alpha$ )
- $Q$  statistic 107–108, 127, 129, 143, 144
- random-effects procedures, *see* meta-analysis: random effects procedures
- randomized controlled trials 89, 115(note 18), 116, 131
- rate ratio, *see* risk ratio
- relative risk, *see* risk ratio
- reliability, *see* measurement; reliability
- replication 43(note 2), 49, 58, 79, 81, 84(note 8), 109, 135, 136
- reporting bias 117
- research synthesis, *see* meta-analysis, narrative review
- reviewer bias 94, 123
- risk difference 7, 13
- risk ratio 7, 8–9, 13
- rival hypotheses, *see* alternative plausible explanations
- $r$  squared, *see* coefficient of determination
- $R$  squared, *see* coefficient of multiple determination
- rugby vs. soccer 31
- robust statistics 28(note 13)
- sample size 47, 63, 67(note 2), 81, 134, 138  
 determination of 57, 60, 61–62  
 measurement error and implications for 66, 67  
 precision and 18, 64–66, 70(note 18), 120  
 rules of thumb and 61  
 statistical power and 56  
 statistical significance and 32
- sampling distribution 20
- sampling error 18, 27(note 12), 67(note 2), 95, 106, 127
- selection bias, *see* availability bias, publication bias, reporting bias
- semipartial correlation, *see* part correlation
- shrinkage 28(note 12)
- significance xiv, 3–4, 5  
 confusion about 4, 5, 24, 25(note 1), 49  
 practical xiii, 3–4, 5, 32, 35, 42, 108, 109, 137  
 statistical xiii, 3–4, 5, 32, 48, 53, 63, 79, 92, 125  
*see-also*  $p$  values, statistical significance testing
- small effects 23, 24, 35–38, 117–119  
 examples of 37, 38, 43(note 7)  
 in elite sports 36  
 that have big consequences 35, 37–38
- Spearman's rank correlation ( $\rho$ ) 11, 13, 15
- SPSS, *see* effect size: SPSS calculations
- squared canonical correlation coefficient 13
- standard deviation 10, 20  
 pooled 10, 26(note 8)  
 weighted and pooled 10, 27(note 9)
- standard error 20, 104, 127, 129
- standard score, *see*  $z$  score
- standardized mean difference 10–11, 15, 21  
*see also* Cohen's  $d$ , Glass's delta, Hedges'  $g$
- Star Wars fans vs. Star Trek fans 33
- statistical power 52–54, 56, 60, 63, 131

Cambridge University Press

978-0-521-19423-5 - The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results

Paul D. Ellis

Index

[More information](#)

Index 173

- effect size and 56, 119, 126  
 how to boost 81–82  
 measurement error and 66, 85(note 14)  
 multivariate analyses, effect on 63–64, 65, 134  
 of published research 73, 75, 76, 83(note 2)  
 precision and 64–66  
 sample size and 52, 56  
 subgroup analyses, effect on 63, 71–72(note 24), 134, 135  
*see also* overpowered tests, power analysis, power calculators, power surveys, underpowered tests  
 statistical significance, *see* significance; statistical  
 statistical significance testing 18, 39, 48, 49, 66, 68(note 4)  
 limitations of 32, 43(note 2), 48, 49, 68(note 5), 115(note 17)  
 misuse of 33, 52, 92, 141  
*see-also* *p* values  
 Super Bowl stock market predictor 51  
 systematic review, *see* meta-analysis
- thresholds for interpreting effect sizes  
 Cohen's thresholds 33, 40–42  
 Rosenthal's thresholds 44(note 13)  
 Pearson's thresholds 44(note 15)
- Titanic* movie 8–9  
*Titanic* survival rates 8
- Tower of Babel bias 120  
 two-tailed test 56, 70(note 15), 81  
 Type I errors 48–50, 51, 54, 56, 79, 94, 133(note 10), 136  
   in meta-analysis 117, 118, 125, 126, 130, 132  
   in published research 55, 80, 82, 84(note 12)  
   statistical power and 56  
 Type II errors 48, 50, 51, 54, 56, 59, 65, 69(note 13), 73, 77, 79, 133(note 10), 135  
   in meta-analysis 117, 124, 130, 132  
   in published research 74–77, 82  
   statistical power and 52, 56, 57, 58  
   *see also* beta-to-alpha ratio
- underpowered tests 52, 82, 92, 124
- variance 104, 106, 107, 131, 133(note 13), 142  
   between studies 129, 144  
   within studies 129, 144  
 vote-counting method 89, 92, 94
- Wilkinson and the Taskforce on Statistical Inference xv, 19, 39, 77, 137  
 winner's curse 132(note 2)
- z* score 104, 105, 143, 149