SUBJECT INDEX

Page references followed by fig indicate an illustrated figure; followed by e indicate an exhibit; followed by t indicate a table; followed by b indicate a box.

A
Accountability: evaluation used to increase transparency and, 799; negotiated, 132
Activities: logic model for increasing graduates entering STEM careers, 76 fig, logic model on the, 64, 65 fig
Adjusted difference of means, 610
Administrative records, 325
Affordable Care Act, 405
Agency records: additional suggestions for obtaining data from, 340–342; as big data source, 825–826; concerns with data accuracy, 330–331; confidentiality and privacy considerations, 336–339; data available only in overly aggregated form, 331–332; data need to be linked across programs and agencies, 335–336; data quality control processes, 339–340; description of, 325; different quality and extensiveness of, 326–327; examples of the content of, 325–326; missing or incomplete data, 327–330; summary of potential problems and possible ways to alleviate them, 328; unknown, different, or changing definitions of data element, 332–335. See also U.S. government programs
Alexa ranking of web traffic, 487
American Association for Public Opinion Research, 375, 486
American Education Research Journal, 691
American Enterprise Institute, 479
American Evaluation Association (AEA): Guiding Principles for Evaluators of, 9, 37, 291, 820; Stafford Hood’s membership and leadership in, 285–286
American Red Cross, 67
American Society for Public Administration, 801
Americans with Disabilities Act, 391
Annie E. Casey Foundation: assessing stability or mobility of unit of analysis, 268; Family Centered Community Change of the, 259; Kid’s Count program of, 811; Making Connections initiative of, 268, 271
Antecedent factors, 66
Anthem Foundation Family Violence Coalitions, 227, 229, 234, 254
Archival records, 325
ArticlesPlus, 477
Assessment instruments: for culturally responsible evaluation (CRE), 294; design
Backcoding” (data editing), 377
Balanced scorecard model, 115
Baltimore’s CitStat, 129
“Bang for the Buck Act,” 801
Bases of power—directions of interest diagram, 51–53
Basic value-added design, 145–146
The Best Evidence Encyclopedia, 689
“Best Evidence for Education (BEE)” [Slavin], 679
Biases:RCT design issue of, 161–162; systematic reviews role of, 673–674
Big data, 825–826
Bing search engine, 476, 674
Biostatistics Department
Brewin-Bradley approach (participant preference design), 164–165
Brookings Institution, 479
Bureau of Community Health Services, 807
Bureau of Labor Statistics’ Longitudinal Employer-Household Dynamics, 272
C
California: Evidence-Based Clearinghouse for Child Welfare of, 690; Health Start program of, 812–813
The Campbell Collaboration: coding categories used in the, 686; Crime and Justice Group of the, 690, 693; efforts to make inclusion standards uniform and scientific by the, 683; literature review using, 13, 475; reviewing studies of intervention effects using, 678, 689; as systematic reviews, meta-analysis, or evaluation synthesis resource, 679, 680, 681, 683, 690, 691, 693, 694–695. See also Literature review
CAQDAS website, 587
Case studies: avoiding common pitfalls, 193–194; definition of, 33; descriptive, 1786–1796; explanatory, 1796; exploratory, 1786; key steps in, 1796–193; pilot-testing, 189; understanding what they are, 177–180
Case study conducting: data collection strategies, 189–190; overview of the, 1796; preparation, 187–189
Case study data analysis: overview of the, 1806; process of the, 190–192
Case study design: choosing single-case or multiple-case designs, 185; defining the research questions, 180–184; determining the unit of analysis, 184–185; overview of the, 1796; selecting cases or sites, 185–187; single-case versus multiple-case designs, 185
Case study example: conceptual framework for food assistance initiative, 1825; on food assistance program modernization initiative, 181–184; research questions addressed and data sources on food assistance initiative, 183–184
Case study report: overview of the, 1806; preparing the, 192–193
Case study steps: analyzing the data, 1806, 190–192; conducting the case study, 1796, 187–190; designing the case study, 1796, 180–187; preparing the report, 1806, 192–193
Casey Foundation. See Annie E. Casey Foundation
Categories (cluster of qualitative codes): clustering qualitative codes into, 565–566; definition of, 564
Cato Institute, 479
Causal inferences: generalizability, 19, 206, 172, 7026, 7046, 7095; internal validity of, 18–196, 161–172, 7026, 7046, 7095; statistical conclusion validity, 20, 7026, 7046, 7095; supporting and ensuring well-founded, 17–18. See also Methodology; Statistics
Causal theory, 140–141
Center for American Program, 479
Center for Mental Health Services: Center for Mental Health Services Housing Initiative of the, 228; McKinney Demonstration Program for Homeless Adults with Serious Mental Illness, 227, 229
Subject Index

Centers of Disease Control and Prevention (CDC): Community Guide initiative of the, 690; Ebola virus epidemic (2014) response by, 413; logic models used by the, 64; National Breast and Cervical Cancer Early Detection Program of the, 132, 133; performance measures of the Division of Sexually Transmitted Diseases Prevention of the, 123, 124;

Champions, 42

Chi-square test, 602, 605–607

Chief Financial Officers Act (1990), 801

Child-care spaces in Ontario, 125;

Child Health Act (1967), 800

Childhood Integrated Longitudinal Data (CHILD), 272

Choice Neighborhoods program (HUD), 259

Cincinnati Public Schools Teacher Evaluation System (TES), 429–430

CIPPEvaluationModelChecklists (Stufflebeam), 767

City of Durham programs, 422–423

"Clean Air Act: Observations on EPA's Cost-Benefit Analysis of its Mercury Control Options" (GAO), 670

Cleanliness Index Monitoring System (Scotland), 425

Cluster analysis, 615

Coalition for Evidence-Based Policy, 327, 679, 683, 689, 829

Cochrane Collaboration: coding categories used in the, 686; efforts to make inclusion standards uniform and scientific by the, 683; literature review using, 13, 475; reviewing studies of intervention effects using, 678, 689; systematic reviews, meta-analysis, or evaluation synthesis using, 679, 680, 681, 683, 690, 691, 694–695. See also Literature review Cochran e Handbook for Systematic Reviews of Interventions, 689

Code of Federal Regulations (U.S. Department of Health and Human Services), 332, 479

Coding: for evaluation syntheses, systematic reviews, and meta-analyses, 684–686; qualitative data, 554–566, 562

Coding qualitative data: categorizing versus, 564–566; clustering into categories, 565–566; description of, 562, 564–565

Cognitive interviewing, 367–368

College and Research Libraries journal, 477

Communication: the audience facet of, 739, 743–745; with gatekeepers (family members of participant), 210; the medium facet of, 739, 745–762; the message facet of, 739, 740–743; multisite evaluations (MSEs) supervision and, 245–246; staff member skills related to, 206–207; telephone calls used for study participant, 210; written contact with student participants, 209–210. See also Evaluation reports

Community change programs: challenges of evaluating, 262–264; data collection sources used for, 272–273; defining interventions for, 260–262; description of, 259–260; examples of challenges and opportunities in evaluating, 276b–278b; guidance for evaluators and practitioners, 264–276. See also Comprehensive community initiatives (CCIs)

Community Development Block Grant (HUD), 276

Community Guide, 690

Community Health Centers program, 807

ComNET2Go initiative, 421–422

Comparison group designs: causal theory and impact evaluation using, 149–154; definition of, 33; description and other names used for, 138–139; failure to provide comparison group during data collection, 709, 714–715; subsidized prekindergarten (pre-K) program outcomes example of, 137–138. See also Randomized controlled trials (RCTs)

Comparison group designs types: basic value-added design, 145–146; fixed-effect designs for longitudinal evaluations, 150–151; interrupted time-series designs, 13, 148–150; matching designs, 151–153; naive design, 142–145; regression-adjusted covariate design, 146–147; regression discontinuity (RD) designs, 153–154; value-added design adjusted for additional covariates, 147–148

Complex adaptive systems, 827

Compliance culture, 14

Compliance reviews, 726, 739

Complier average causal effect (CACE) analysis, 163

Comprehensive community initiatives (CCIs), 260, 261, 364; defining possible interventions of, 261; description of, 260; theories of change used by, 263. See also Community change programs

CompStat, 129, 801

Computer-assisted telephone interviewing (CATI or CAPI), 370, 371, 373

Computerized Neighborhood Environment Tracking (ComNET), 418–421

Concept plan (or purchase plan): description of, 767–768; key...
844

Subject Index

Cost-benefit analysis steps: step 1: set the framework for analysis, 638b–639b; step 2: decide whose costs and benefits should be recognized, 640–642; step 3: identify and categorize costs and benefits, 642–643b; step 4: project cost and benefits over the life of the program, 645–646b; step 5: monetizing (putting a dollar value on) costs, 646–651b; step 6: quantify (for CEA) and monetize (for CBA) benefits, 651–659b; step 7: discount costs and benefits to obtain present values, 659–661b; step 8: compute cost-effectiveness ratio (for CEA) or net present value (for CBA), 661–664, 663b; step 9: preform sensitivity analysis, 664–666, 665b; step 10: make a recommendation, 666–669.

Cost-benefit studies (CEA): deciding on CBA or, 638b–639b; definition of, 35, 63b; for multisite adult drug court evaluation, 446b–447b; selected applications and critiques of, 669b–671b; steps in, 637–669. See also Dropout prevention program

Cronbach’s alpha, 17

Cross-tabs: description of, 598; with two ordinal variables, 599.

The crowd (stakeholder), 49

Cuban missile crisis case study (1971), 185

Cultural competence (cultural humility): demonstrated toward stakeholders, 40–41; growing evaluation importance of, 13–14; related to study participants, 219–220

Culturally responsible evaluation (CRE): case applications of, 296–302; core literature on key characteristics of, 288b–290b; Culture Checklist to use for, 306, 307b–309b; data analysis for, 295; data collection for, 294–295; description of, 282–283; designing the evaluation, 293–294; disseminating and using the results, 295–296; engaging stakeholders, 291–292; framework of, 287b; framing the right questions, 292–293; identifying purpose and intent of, 292; implications for the profession, 302–303–306; introduction to, 281–282; justifications of validity under, 304b–305b; Key Evaluation Checklist (KEC) used for, 302, 306; Māori evaluation, 310; pioneers in the foundations of, 283–287; by pre-Brown African American evaluators, 306, 310; preparing for the, 290–291; responsibility as a core principle of, 306, 309–310; selecting and adapting instrumentation, 294

Culturally responsive school-based evaluation (Southwestern U.S.), 297–299

Customer satisfaction measure: description of the, 113; disability adjudicators’ claims processing performance, 120

elements of a, 769–770; shaping a feasible, 770–773; tips on creating a feasible and approved, 768b–769b

Confidence interval, 605

Confidentiality issues: agency records and, 336–339; Education Commission of the States (1991) on meaning of, 819; evaluation stories and, 549–550; Family Educational Rights and Privacy Act (FERPA), 272, 337; field visit sample introduction and statement of, 468b; Health Insurance Portability and Accountability Act (HIPAA) on, 211, 212, 237, 819; multisite evaluations (MSEs) and, 247–248; security of the resulting data to protect, 818. See also Ethical issues; Information; Legal issues; Privacy issues

Congressional Research Service (CRS) Report, 478

CONSORT statements (health care research), 173

Context sellers (stakeholder), 49

Contingency tables: description of, 598; with two ordinal variables, 599.

Contracting. See Evaluation contracting

CoreLogic, 273

Corporation for National and Community Service, 327

Corporation for Supportive Housing Employment Initiative, 227b, 228

Cost-benefit analysis (CBA): deciding on CBE or, 638b–639b; definition of, 63b; formula for cost-effectiveness ratio, 636–637; selected applications and critiques of, 669b–671b; steps in, 637–669; units of effectiveness, 636–637. See also Dropout prevention program
Subject Index

D

Data: “backcoding” (editing), 377; big data, 825–826; case study program, 190; challenges to anticipate related to use of, 31; evaluation, 30–32; inventory of available, 271–273; multisite evaluations (MSEs) management of, 246–248; pitfalls related to inadequately support of conclusions with specific, 722; qualitative, 12, 16–17, 543, 561–566, 562; quantitative, 12, 16, 17, 271–273, 563–564. See also Performance data

Data analysis: case study, 180; 190–192; culturally responsible evaluation (CRE), 295; of focus group interview data, 525–532; matching evaluation questions to, 271–291; of performance data, 127–129; quantitative, 245–253, 508; reliability issue of, 16–17; selecting the right unit for evaluating community change programs, 265–266; software packages used for, 246, 254, 607, 613, 616, 617; survey, 376–377; tips for presenting, 617b–621b; triangulation used for, 191. See also Qualitative data analysis; Statistics; Unit of analysis

Data checks, 376

Data collection: case study, 187–188, 189–190; case study pitfalls of collecting too much information during, 181; choose appropriate measures, 14–16; choose reliable ways to obtain the chosen measures, 16–17; community change program, 271–273; culturally responsible evaluation (CRE), 294–295; field visit protocol on, 468b–470; guide for review of evaluability assessment (EA), 936, 94b; identify contextual elements that may affect evaluation conduct and use, 13–14; interviews used for, 189, 371–372, 463–467, 492–504, 509–533; IRB approval of new methods of, 274; location of, 292; logic model construction stage of, 71–72; matching evaluation questions to, 271–291; of multisite evaluations (MSEs), 235–238, 243; participant motivation and modes of, 201–202; produce the methodological rigor needed to support credible findings, 14; qualitative, 271–273, 492; role playing, 390–392; select programs to evaluate, 9–10; select the type of evaluation, 10–12; surveys used for, 272, 344–382. See also Evaluation data; Information

Data collection pitfalls: after data collection, 716–723; before data collection begins, 703–708; during data collection, 708–716

Data collection rapid feedback assessment (TFE), 100

Data collector training, 704b, 708

Data sources: agency records, 325–343; big data, 825–826; case study of food assistance initiative, 183–184; community change programs, 272–273; of performance data, 115–116; U.S. Department of Education on problems with whose costs and benefits should be recognized, 327

Data systems: creation and management of a, 273–274; description of, 273

Data visualization, 826

DataQuick, 273

Decennial Census, 272

Delphi method, 509

Demographix, 487

Demonstrated evidence-based interventions (DEBs), 828

Department of Veterans Affairs, 540

Depth interviewing, 492

Descriptive case study, 178b–179b

Descriptive methods: Breast Services Enhancement Program (BSEP) application of, 574; compared to other methods, 567; matrix displays example of, 572–574, 575; qualitative data analysis using the, 566, 572; when to use, 575–576

Descriptive statistics: bivariate, 598–599; conditional, 596, 598–599; overview of, 597; univariate, 597

Direct observation, 190

Disability adjudicators’ claims processing performance indicators, 129

Discriminant function analysis, 615


Disparities Research for Change, 203

Disseminating, See Evaluation reports

Document reviews (case study), 190

DoView software, 81b

Dropout Prevention Program:

step 1: set the framework for the analysis, 640; step 2: decide whose costs and benefits should be recognized, 641–642; step 3: identify and categorize costs and benefits, 643–645; step 4: project cost and benefits over the life of the program, 646; step 5: monetizing (putting a dollar value on) costs, 649–651; step 6 (for CBA): quantify for CBA, 657–659; step 6 (for CEA): quantify for CEA, 652–657; step 7: discount costs and benefits to obtain present values, 661; step 8: computer cost-effectiveness ratio (for
Subject Index

CEA) or net present value (for CBA), 663–664; step 9: perform sensitivity analysis, 665–666, 665f; step 10: make a recommendation, 668–669. See also Cost-benefit analysis (CBA); Cost-benefit studies (CEA)

Dun and Bradstreet, 273

DuPont Corporation, 337

E

Early Childhood Longitudinal Studies, 150

Econometric models: classical content analysis example of, 567–569; compared to other methods, 567f; qualitative data analysis using the, 566; United Nations Global Pulse Lab application of, 569–570f; when to use, 570–571

Environmental Protection Act (1990) [Scotland], 423

ERIC database, 691

Evaluability assessment (EA): process: step 1: involve intended users and other key stakeholders, 92; step 2: clarify the program design, 92–93; step 3: explore program reality, 93–94; step 4: assess the plausibility off the program, 94; step 5: reach agreement on any needed changes in program design or implementation, 95; step 6: reach agreement on evaluation focus or intended use, 95–96

Evaluation Assessment and Policy Connections Unit (University of North Carolina), 489

Evaluation conduct: biases that may affect, 161–162, 673–674; identifying contextual elements that may affect, 13–14; overly intrusive data collection procedures that change participant behaviors (Hawthorne effect) on, 709, 712–713

Evaluation contracting: assuring product quality and usefulness, 794–796; creating a feasible and approved concept plan for, 796–773; five principles for successful, 760f; the goal of, 765; limited competition practice of, 773; monitoring interim progress, 789–794; overlap of information technology (IT) services with, 768; overview of, 765–767; selecting a well-qualified contractor, 780–789. See also RFP (request for proposal)

Evaluation contractors: conducting final negotiations with, 789; description of the, 780–781; process of selecting the, 782–789; reviewing proposals of, 781–782; tips on selecting a well-qualified, 781f

Evaluation decisions: approve the most appropriate evaluation

Effect size: analysis strategy for, 687–688; computing, coding, and estimating their variances, 686–687; description of, 686; estimating, 686–687; under-emphasis on significance of, 716–718

Effectiveness evaluation, 159

Efficacy evaluation, 159

Efforts to Outcomes, Salesforce, and Community Techknowledge (or CTK) [Social Solutions], 273

Eight Year Study, 286

Electronic reports: description of, 760–761; tips on producing reliable, 761f

Elementary and Secondary Education Act (1965), 799–800

Elite interviewing, 492

Emotional Intelligence (Goleman), 535

Enhancing Professional Practice: A Framework for Teaching (Danielson), 430

Enumerative methods: classical content analysis example of, 567–569; compared to other methods, 567f; qualitative data analysis using the, 566; United Nations Global Pulse Lab application of, 569–570f; when to use, 570–571

Ethical standards and issues: challenge of maintaining evaluation standards and ethics, 818–820; GAO’s Yellow Book standards, 820; Guiding Principles for Evaluators (AEA), 9, 37, 291, 820; informed consent, 819; IRBs to ensure ethical treatment of human subjects, 203–204, 210, 212, 222, 229–233, 274, 337, 465, 819–820; Joint Committee on Standards for Educational Evaluation (1994), 37; Joint Committee on Standards for Educational Evaluation (2010), 9, 820; privacy as, 211, 212, 247–248, 272, 336–339; role playing studies and, 406–407. See also Confidentiality issues

Ethnographic interview, 492

Evaluation Assessment and Policy Connections Unit (University of North Carolina), 489

Evaluation conduct: biases that may affect, 161–162, 673–674; identifying contextual elements that may affect, 13–14; overly intrusive data collection procedures that change participant behaviors (Hawthorne effect) on, 709, 712–713

Evaluation contracting: assuring product quality and usefulness, 794–796; creating a feasible and approved concept plan for, 796–773; five principles for successful, 760f; the goal of, 765; limited competition practice of, 773; monitoring interim progress, 789–794; overlap of information technology (IT) services with, 768; overview of, 765–767; selecting a well-qualified contractor, 780–789. See also RFP (request for proposal)

Evaluation contractors: conducting final negotiations with, 789; description of the, 780–781; process of selecting the, 782–789; reviewing proposals of, 781–782; tips on selecting a well-qualified, 781f

Evaluation decisions: approve the most appropriate evaluation

Evaluation Assessment and Policy Connections Unit (University of North Carolina), 489

Evaluation conduct: biases that may affect, 161–162, 673–674; identifying contextual elements that may affect, 13–14; overly intrusive data collection procedures that change participant behaviors (Hawthorne effect) on, 709, 712–713

Evaluation contracting: assuring product quality and usefulness, 794–796; creating a feasible and approved concept plan for, 796–773; five principles for successful, 760f; the goal of, 765; limited competition practice of, 773; monitoring interim progress, 789–794; overlap of information technology (IT) services with, 768; overview of, 765–767; selecting a well-qualified contractor, 780–789. See also RFP (request for proposal)

Evaluation contractors: conducting final negotiations with, 789; description of the, 780–781; process of selecting the, 782–789; reviewing proposals of, 781–782; tips on selecting a well-qualified, 781f

Evaluation decisions: approve the most appropriate evaluation
Subject Index

method, 274–276; choose appropriate measures, 14–16; choose reliable ways to obtain the choose measures, 16–17; identify contextual elements that may affect evaluation conduct and use, 13–14; produce the methodological rigor needed to support credible findings, 14; Pyramid of Strength to provide credible findings of, 14, 15; select programs to evaluate, 9–10; select the type of evaluation, 10–12; survey planning and related, 346–357; for trained observer ratings, 417. See also Program evaluations

Evaluation designs: culturally responsible evaluation (CRE), 295–294; definition of, 33; evaluability assessment, 11; for field studies, 450–451; for formative evaluation, 11; implementation evaluation, 11; likely to generate false positives or false negatives, 603; matching evaluation questions to, 274–276; multisite evaluations (MSEs), 233–234; online survey, 484–486; process studies, 11; regression discontinuity (RD), 34, 153–154; sample design matrix to use for, 267; single-case study, 185; stepped wedge, 165, 166–167; U.S. Government Accountability Office (GAO) approach to, 226–258; value-added, 145–148; waiting list, 165–167

Evaluation findings:

generalizability of, 19, 20, 172, 702, 704, 709; methodology rigor to support credible, 14; of multisite evaluations (MSEs), 255; Pyramid of Strength to provide credible, 14, 15; supporting causal inferences of, 17–21; tips on using evaluation data and, 326. See also Evaluation reports

Evaluation information: effective application of, 30–32; tips on using findings and, 326. See also Information

Evaluation mandates, 827–828

Evaluation planning: the evaluation processes, 21–22; evaluation stakeholder engagement planning, 292–293; defining case study research questions, 180–184; evaluation story addressing the, 546; fieldwork studies, 451–453; focus group interview, 516–518; on food assistance modernization initiative case study, 183–184; matching data collection to, 274–276, 30; matching designs to, 274–276; matching them to client’s information needs, 25–30; revising them as you learn more during evaluation processes, 22; See also Survey questions

Evaluation reports: the audience facet of, 739; 740–743; performance data, 127–129; pitfalls related to poor presentation, 722–723; putting program evaluation on the Internet form of, 489–490; reporting statistical result to high-level public officials, 621–622; of semi-structured interviews (SSIs), 504; survey, 378–381; of systematic reviews, 688–689; trained observer ratings, 435–438; writing style and layout, 752, 756–762b. See also Communication; Evaluation findings; Foster Care Parent Recruitment Report; Outcomes

Evaluation sponsors: approval of contracting concept plan by, 767–773; definition of, 42; deliverable projects submitted to, 776; developing well-defined RFP (request for proposal), 773–780; monitoring interim progress of evaluation contracting, 789–794

Evaluation stakeholder engagement: in culturally responsible evaluation (CRE), 295–292; evaluation stakeholder engagement planning matrix, 56–57; Evaluation stakeholder types: key stakeholders, 37; primary intended users subset of key stakeholders, 37–41; specific stakeholders, 37

Evaluation stakeholders: acknowledging the importance of, 36; dealing with power differentials, 48–53; definition of, 33, 37; demonstrating cultural competence toward, 40–41; engaging, 55–58; evaluability assessment (EA) role of, 92, 93; failure to secure input on appropriate evaluation criteria from, 704; 705–706; four categories of, 49;
identification and analysis techniques used with, 41–48; managing challenges of turbulent and uncertain environments of, 58–59; perceptions of problems by, 10; understanding who they are, 37–38; verifying program logic model with, 82–85; vet recommendation ideas with, 731–732. See also Primary intended users

Evaluation stories: challenges in using and managing, 554–555; comparing strategies for using, 544; concept of truth in, 551–552; confidentiality issue of, 549–550; decide on the evaluation question or topic of the, 546; definition of a, 537–538; examining the power of, 535–536; how they can be used in evaluation studies, 539–543; how they enrich evaluations, 536–537; key characteristics of, 538; “medical mistakes” example of a, 538; overview of critical steps in using, 543, 546–552; sampling strategy to use with, 546; strategies of expert storytellers for presenting, 532–554

Evaluation synthesis: contributing to a meta-analysis, systematic review, and, 680; definition of, 89, 90; overview of the, 104; purpose and cost of, 90; resources employed to do a good job on, 689–682; steps for producing a, 680–681; value added and usefulness of systematic reviews and, 692–693; when to use, 106

Evaluation types: case studies, 33, 177–195; evaluability assessment (EA), 88, 90; evaluating synthesis, 89, 90; 104; formative, 11; impact, 140–154, 447, 448; implementation, 11, 454–456; rapid feedback evaluation (RFE), 89, 90, 98–103; small-sample studies, 89, 90; 104–105; summative, 11; theory-based evaluation (TBE), 54

Evaluator tasks: anticipate turnover of intended user, 41; demonstrate cultural sensitivity and competence, 40–41; develop facilitation skills, 38; facilitate high-quality interactions, 40; find and train evaluation information users, 38–39; find tipping point connections, 39–40; nurture interest in evaluation, 40

Evaluators: acknowledging the importance of stakeholders, 36; challenge of selecting and training, 817–818; cultural competence (cultural humility) of, 13–14; danger of compliance culture to objectivity of, 14; dealing with power differentials, 48–53; focus group interview responsibilities of, 512; guidance for evaluating community change programs, 264–276; making recommendations for program improvement, 725–728; managing challenges of turbulent and uncertain environments, 58–59; pre-Brown African American, 506, 310; tips for the multisite evaluations (MSEs), 255–256

Evidence-Based Clearinghouse for Child Welfare (California), 490

Evidence-based policy, 13

Evidence-based practice: debate over how much evidence is required for, 13; demonstrated evidence-based interventions (DEBs), 82; identifying high-quality evidence in intervention effects studies, 676, 678; increased demand for rigorous, 828–830

Evidence for Policy and Practice Information and Coordinating Centre (EPPI Centre) [University of London], 679

Excel, 585, 638

Executive Orders of the President, 479

Executive summary: description of, 746, 749–750; Foster Parent Recruitment Report example of, 754–755; tips on writing the, 751

Experimental design: comparison group designs, 137–155; definition of, 33; interrupted time-series design, 33; quasi-experimental, 34, 275; randomized controlled trials (RCTs), 13, 33, 157, 158–175; regression discontinuity design, 34

Experts: definition of, 15; evaluation role of, 15–16

Explanatory case study, 178

Explanatory methods: compared to other methods, 567; example of truth table, 582, 583; qualitative comparative analysis (QCA) example of, 581–582; qualitative data analysis using the, 566, 580–581; Swiss Environmental Impact Assessment (EIA) application of, 582, 584; when to use, 584

Exploratory case study, 178

Exploratory evaluation: evaluable assessment (EA) approach to, 88, 90; 106; evaluation synthesis approach to, 89, 90; 104, 106; rapid feedback evaluation (RFE) approach to, 89, 90; 98–103, 106; selecting which approach to use for, 105–106; small-sample studies approach to, 89, 90; 104–105, 106

External benchmarking, 124–126

External validity: generalizability or transferability of, 19, 20,

Subject Index
subject index

702b, 704t, 709t; randomized controlled trials (RCTs), 172
Extraordinary Knowing (Mayer), 542
EZ-Survey, 369

F

Face-to-face surveys: collecting data using, 351t, 353t, 370; computer-assisted personal interviewing (CAPI), 370, 371, 373, 374, 376; contact record sheet for, 374e
Face validity, 15
Facilitation skills development, 38
Fair Housing Act, 389, 391
False negative (type II error), 450, 460–461
False positive (type I error), 450, 460–461
Facilitator activities, 504
Facilitator role, 503
Facilitator responsibilities, 503
Facilitator training, 504
Facilitation skills development, 38
Facilitation skills development, 38
Field study frameworks: data collection guided by the, 451–456; implementation, 454–456; logic models, 453; research questions, 451–453
Field study sites: basis for site selection, 459–461; number of, 457–459; site packets for, 466–467; types and scope of instruments used at the, 461; unit of analysis issue of, 456–457
Field visit protocol: on-site procedures, 467–470; overview of, 461–462; previsit preparations, 462–467
Fieldwork models: program evaluation, 448–450; program management, 448
Fieldwork models: program evaluation, 448–450; program management, 448
Field of study objectives: for impact evaluation of aftercare program for substance abusers, 448t; for multisite adult drug court evaluation, 446b–447b; overview of, 446t; for program outcome evaluation of a child support enforcement collections system, 447t for welfare-to-work programs in New York City, 447t
Field study frameworks: data collection guided by the, 451–456; implementation, 454–456; logic models, 453; research questions, 451–453
Field study sites: basis for site selection, 459–461; number of, 457–459; site packets for, 466–467; types and scope of instruments used at the, 461; unit of analysis issue of, 456–457
Field visit protocol: on-site procedures, 467–470; overview of, 461–462; previsit preparations, 462–467
Fieldwork models: program evaluation, 448–450; program management, 448
Findings. See Evaluation findings
Fixed-effect designs for process studies, 11
Flesch-Kincaid assessment, 209
Focus group interviews: addressing challenges in, 583; analysis of, 525–532t; characteristics of, 509–512; developing questions for, 516–518; evaluator responsibilities during, 512; moderating, 522–525; planning, 513–515; recruiting for, 519–521
Focus groups: description of, 506; examples of uses of, 507–509
Focused interview, 492
Food assistance modernization initiative case study: conceptual framework for, 182f; overview of the, 180–184t; research questions addressed and data sources for, 183t–184t
Food Policy Institute (Rutgers University), 489
Formative evaluation: focus group interview used for, 518; issues to consider for using, 11f; varieties of designs available for, 11
Formative evaluation designs: evaluability assessment, 11; implementation evaluation, 11; process studies, 11
Forsight platform, 569
Foster Care Parent Recruitment Report: background of the, 746–747t; executive summary of the, 754t–755t; killer paragraph for the, 749t; Mom Test summary for, 747t; slide presentation for the, 763t; topic outline for the, 750t–751t; two examples of using graphics in the, 757t–758f; See also Evaluation reports; The medium formats
FreeOnlineSurveys, 487
From State Hospital to Psychiatric Center: The Implementation of Planned Organizational Change (Levine), 179t
Fund for Modern Courts, 433, 434t

G

Gatekeepers (family members of participants), 210
Generalizability: description of, 19; questions to ask strengthen the finding, 20t; of randomized controlled trials (RCTs), 172; as touchstone of methodological integrity, 702b, 704t, 709t
GFKKnowledge Panel, 353
GIS (geographical information systems), 826
Goals: determining evaluation purpose and, 53–55; evaluation contracting, 765; program
Subject Index

improvement as program evaluation, 8; strategy maps linked to program, 115; U.S. Department of Health and Human Services’ “promote safety,” 113, 115
Google Scholar, 476–477
Google search engine, 476, 674
Government. See U.S. government programs
Government Performance and Results Act (GPRA), 770, 801, 821
GPRA Modernization Act (2010), 802–804, 821
Graphics: Foster Parent Recruitment Report examples of using, 757–758
fig; tips on using large, 757; types on using, 756
“Gray literature,” 476, 478
GreenBook, 486
Guiding Principles for Evaluators (AEA), 9, 37, 291, 820

H
Harlem Valley Psychiatric Center (New York State), 179
Harris Interactive, 353
Harvard Family Research Project (Harvard University), 489
Havens Reality Corp. v. Coleman, 407
Hawthorne effect, 709; 712–713
Head Start: propensity score matching used to study, 152; Westinghouse study of, 151
Health Insurance Portability and Accountability Act (HIPAA), 211, 212, 272, 337, 819
Healthy People (2010), 813
Healthy Schools evaluation, 226, 227
Healthy Start (California), 812–813
Heritage Foundation, 479
Subject Index

strengthening abilities to meet increasing demand for, 8
Information technology (IT): increasing trend of program evaluation use of, 824–825; overlap of evaluation contracting services with services of, 768
Informed consent, 819
InfoUSA, 273
Inputs (State Board of Nursing Program Logic), 114
Institute for Education Sciences, 681, 682
Institute for Social Research (University of Michigan), 490
Institute of Public Policy (George Washington University), 489
Institutional review boards (IRBs): approval of new data collection methods by, 274; field visits and role of, 465; multisite evaluations (MSEs) oversite by, 232–233; protecting rights of study participants, 203–204, 210, 212, 222, 337, 819–820
Intention to treat (ITT) analysis, 163, 165
Interagency collaborative capacity (ICC), 812–813
Intermediate outcomes: logic model for increasing graduates entering STEM careers, 76; logic model on, 65; State Board of Nursing Program, 114
Internal validity: causal inferences and, 18–19; description of, 18; of randomized controlled trials (RCTs), 161–172; as touchstone of methodological integrity, 702, 704, 709
International Association for Public Participation, 55
International Federation of the Red Cross and Red Crescent Societies, 682
International Initiative for Impact Evaluation, 676
The Internet: Alexa ranking of web traffic on, 487; ArticlesPlus, 477; conducting surveys on, 351; 353, 369–370, 479–489; literature reviews using, 474–479; PAIS International, 477; ProQuest Central, 477; putting your program evaluation on, 489–490. See also Technology Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method (Dillman, Smyth, and Christian), 354
Internet surveys, 351; 353, 369–370
Interrupted time-series design: definition of, 33; impact evaluations using, 148–150
Interval statistics: description of, 596; matching statistical techniques to analytical objectives, 614. See also Statistics
Intervention effects studies: computing, coding, estimating, and analysis of effect size variances, 686–688, 716–718; going beyond the flaws in conventional literature reviews, 678; identifying high-quality evidence, 676, 678; multiple evaluations versus a single evaluation, 674–675; pitfall of using inadequate indicators of program effects, 704, 707–708; processing effects on prevalence: first effects, 677; systematic reviews, meta-analyses, and evaluation synthesis of, 673, 674–675, 680–695
Interviews: case study, 189; field protocol for on-site, 467; field visit, 463–464; focus group, 509–533; semi-structured, 492–504; training survey interviewers, 371–372

J
Jerry Lee Foundation, 690
Job Corps, 800, 808

K
Keep America Beautiful (KAB), 417–418
Keep Scotland Beautiful, 425
Key Evaluation Checklist (KEC), 302, 306
Key stakeholders, 37
KeySurvey, 487
Kids Count (Casey Foundation), 811
Killer paragraph: description of, 746, 748; Foster Parent Recruitment Report example of, 749; tips on writing the, 748
KPMG Foundation, 669–671

L
Layout. See Writing style and layout
LEAMS (A Local Environmental Audit and Management System) [Scotland], 424–425
Legal issues: compliance culture, 14; IRBs to ensure ethical treatment of human subjects, 203–204, 210, 212, 222, 232–235, 274, 337, 465, 819–820; role playing studies and, 406–407. See also Confidentiality issues; U.S. legislation
LIFT Program, 267
Limited competition contracting, 773
LineService, 487
Logic model construction: stage 1: collecting the relevant information, 71–72; stage 2: clearly defining the problem and its context, 72–74; stage 3: defining the elements of the program in a table: early sense making, 74–77; stage 4: drawing the logic model to reveal the program’s theory of change, 77–81; stage 5: verifying the program logic model with stakeholders, 82–85.

Logic model examples: for increasing middle school STEM programs, 75, 76/fig–79/fig, 83–84/fig, 85; two levels of detail for increasing graduates entering STEM careers, 79/fig.

Logic models (or program logic model): advantages for conducting evaluations, 62–65; benefits of using, 68/fig; building a, 70–83, 85; descriptions of the elements of the, 64, 66; definition of, 33; for describing program theory’s theory of change, 68–70; diagram of a basic, 65/fig; example for increasing middle school STEM career graduates, 75, 76/fig–79/fig, 83–84/fig, 85; fieldwork studies, 453; State Board of Nursing Program, 114/fig; tips to consider before starting your, 70b–71b; the utility of, 67–68.

Long-term outcomes: logic model for increasing graduates entering STEM careers, 76/fig; logic model on, 65/fig, 66; State Board of Nursing Program, 114/fig.

Low-Income Housing Tax Credit (U.S. Treasury), 273, 276.

M

Mail surveys, 351t, 352, 366, 368.

Making Connections initiative (Casey Foundation), 268, 271.

Managers: failure to secure input on appropriate evaluation criteria from, 704t, 705–706; failure to clarify expectations about what can be learned from evaluation, 704t, 706; guide for evaluability assessment (EA) review of interviews with, 93t, 94t; “How Are We Doing?” sessions held by, 821; performance incentives for government program, 810t. See also Programs; Staff.

Mandatory evaluation trend, 827–828.

Māori evaluation, 310.

MARTA (Metropolitan Atlanta Rapid Transit Authority), 121–122/fig.

Maryland’s Systems Reform Initiative, 813.

Matching designs, 151–153.


Maturation (participant), 709t, 714.
Subject Index

Meta-analyses: contributing to a, 680; definition o, 673; of multiple evaluations, 674–675; resources employed to do a good job on, 689–692; steps for producing a, 680–689; value added and usefulness of systematic reviews and, 692–693
Methodology: choose appropriate measures, 14–16; choose reliable way to obtain chosen measures, 16–17; the evaluation report message on, 742–743; overview of the reliability issue of, 16–17; overview of the validity issue of, 15–16, 18–19; Pyramid of Strength to provide credible findings, 14, 15;fg supporting credible findings with rigorous, 14; tips on enhancing reliability, 17. See also Causal inferences; specific methodology
Methodology integrity: touchstones: credibility, 14–17; 702b, 704b, 717b; generalizability or external validity (transferability), 19, 20, 172, 702b, 704b, 717b; internal validity (confirmability), 18–19b, 65fg, 161–172, 702b, 704b, 717b; measurement validity (trustworthiness and authenticity), 702b, 704b, 717b; pitfalls occurring before data collection begins, corresponding to, 704b, 717b; reliability (auditability), 16–17b, 703b, 704b, 717b; statistical conclusion validity, 20, 702b, 703b, 704b, 717b. See also Pitfalls in evaluations
Metropolitan Atlanta Rapid Transit Authority (MARTA), 121–122fg
Microsoft Excel, 585, 638
Microsoft Office Online, 209
Milbank Fund, 694
Mixed-method approaches, 12
Mixed-mode surveys, 333–354
Moderating focus group interviews: moderator responsibilities for, 523–525; moderator skills for, 522–523
The Mom Test, 740–741, 742, 764b
The Mom Test summary: description of the, 746; Foster Parent Recruitment Report example of, 747b; as part of the formula for success, 764b
Monitoring: focus group interview used for agency evaluation and, 508–509; multisite evaluations (MSEs), 241–242; progress of evaluation contracting, 789–794; progress of recruitment and retention of participants, 216–219; relationship between performance evaluation and, 822–824; role playing used for, 386–387
MonkeyLearn, 585
Multiple-case case study design, 185
Multiple regression model, 609–611
Multisite evaluations (MSEs): advantages and disadvantages of, 230; assessing interventions, 238–240; data collection, 235–238, 243; data management during, 246–248; defining the, 226–229; determining the design, 233–234; examples of, 226, 227; 228–229; Institutional Review Board (IRB) overseeing process of, 232–233; laying the foundation for an, 231–233; monitoring implementation of, 241–242; overview of, 225; qualitative data analysis strategies for, 253–255; quality control in, 242–246; quantitative analysis strategies for, 248–253; sampling sites, 234–235; telling the story when reporting findings, 255; tips for the MSE evaluator, 255–256; training and booster sessions for, 243–244
Naive design, 142–145
National Academy of Public Administration, 801, 808
National Academy of Sciences (NAS), 693
National Breast and Cervical Cancer Early Detective Program, 132, 133fg
National Center for Education Statistics Common Core, 272
National Institute of Corrections, 476
National Neighborhood Indicator Partnership, 272
National Opinion Research Center, 213
National Public Radio, 339
National Student Clearinghouse, 273
Negotiated accountability, 132
Neighborhood Improvement Initiative (Hewlett Foundation), 277b–2788
New Directions for Evaluation, 286
New Hampshire-Dartmouth Psychiatric Research Center, 244
New Jersey Work Incentive experiment, 719, 720–721
New Markets Tax Credit program (U.S. Treasury), 262, 271, 273, 276
New York American Marketing Association, 486
New York City: field studies objectives for welfare-to-work program in, 447b; Mayor’s Office of Operations in, 423; New York City School System, 130; New York Police Department’s CompStat, 129, 801; New Yorkers for Parks (NY4P) of, 427–428
Subject Index

854

The New York Times, 535, 693
No Child Left Behind Act (2002), 118, 298, 802
Nobody Knows My Name (Hood), 286
Nominal statistics: description of, 596; matching statistical techniques to analytical objectives, 614f; statistical hypothesis testing, 601–611. See also Statistics
Null hypothesis, 602, 604, 606
Ordinal statistics: contingency table with two ordinal variables, 598–599f; contingency tables
Objective Structured Clinical Examination (OSCE), 433, 435
Objectives (field studies), 446b–448b
Observation: case study, 190; decisions about ratings and sampling, 417; examining the quality of data obtained through, 412–413; examples of trained observer programs for, 417–426; observing and rating interactions, 428–442; operating trained observer initiatives, 443b; outsiders running trained observer programs, 426–428; uses for trained observer ratings, 413; what you will need to start, 416; when to use a trained observer method for, 414b–416
Office of Inspector General (OIG), 746
Office of Management and Budget (OMB), 203, 327, 802, 803
Ohio Department of Transportation (DOT) indices, 127–128f
Ontario child-care spaces, 125fg
Ordinal statistics: contingency table with two ordinal variables, 598–599f; contingency tables
Panel Study of Income Dynamics, 150
Panel Study of Income Dynamics, 150
Performance measures; Outputs: definition of, 34; failure to take into account contextual factors that affect, 709f; 715, 717f; focus and trends for performance measures on, 134–135f; logic model for increasing graduates entering STEM careers, 76f/g; logic model on the, 65f/g; 66; as performance measure, 110–111; short-term, 76f/g, 114f/g; State Board of Nursing Program Logic on, 114f/g; tips for results-driven approach to identifying relevant, 135f. See also Evaluation reports; Performance measures; Programs
Outliers (Gladwell), 535
Outliers (statistical), 616
Outline (report): description of the, 746, 748, 749; Foster Care Parents Recruitment report, 750b–751b
Outputs: definition of, 34; disability adjudicators’ claims processing performance measure of, 120f; logic model for increasing graduates entering STEM careers, 76f/g; logic model on the, 64, 65f/g, 66; as performance measure, 111; State Board of Nursing Program Logic on, 114f/g. See also Performance measures
PAIS International, 477
Panel studies: Early Childhood Longitudinal Studies, 150;
Panel Study of Income Dynamics, 150
Panel Study of Income Dynamics, 150
Panel Study of Income Dynamics, 150
Paperwork Reduction Act (PRA), 203
Participant preference design (Brewin-Bradley approach), 164–165
Participants: cultural considerations related to, 219–220; data collection modes used with, 201–202; defining the target population of, 199–201; failure to account or natural maturation among, 709f, 714; failure to draw sampling representative of program, 709f, 713; failure to take into account degree of difficulty of helping program, 709f, 715–716f; frequently asked questions (FAQs) provided to, 214–215; gaining cooperation of, 212–216; in a health care setting, 210–212; implementing recruitment and retention of, 209–216; informed consent of, 819; IRB role in rights and welfare of, 203–204, 210, 212, 222, 232, 274, 337, 465, 819–820; monitoring recruitment and retention progress of, 216–219; motivation of, 201; overly intrusive data collection procedures that change behaviors of, 709f, 712–713; planning or recruitment and retention of, 198–209; pretesting, 202–203; pros and cons of incentives for, 215–216; reasons for refusing to participate and potential rebuttal approaches, 213f; recruiting focus group interview, 519–521; recruitment and retention challenges and possible solutions, 220, 221f; recruitment and retention overview, 197–198; social norms for, 213f; staff role in

P
Subject Index

recruitment and retention of, 204–208; written contact with, 209–210

“Pay for Success” (PFS) programs, 828

Performance data: broken out by different types of cases, 125–126; challenges to anticipate related to use of, 314; converting to information, 121–126; evaluation synthesis used to summarize program, 89, 90, 104; generated by performance measurement systems, 108–135; logic model on middle school STEM program, 83, 84/fg, 85; Ohio Department of Transportation (DOT) indices, 127–128; presenting and analyzing, 127–129; red flag that can be raised by, 110; relationship of Z model to, 79/fg, 83; sources of, 115–116. See also Data

Performance incentives: creating incentives for government program, 809–811; financial, 810, 811; for managers, staff, and organizational, 810f

Performance measurement systems: examining how to design and implement, 108; performance partnerships use information from, 812; tracking program or agency performance over time, 109–110; types of measures tracked by, 110–113

Performance measures: actual performance versus target performance, 122, 123f; comparisons among units, 122–123, 124/fg; cost-effectiveness, 111, 118–119, 120f; criteria for good, 116–119; current challenges to, 129–134; customer satisfaction, 113, 120f; description of practice of, 108; developing agreement on government indicators for, 811–812; efficiency, 112; identifying, operationalizing, and assessing, 113–119; implementing in networked environments, 131–134; increasing demands for, 108–109; program evaluation and role of, 109–110; relationship between evaluation and monitoring, 822–824; review of disability adjudicators’ claims processing, 120f; risk-adjusted, 810; service quality, 112, 119, 120f; small-sample studies used for vetting, 89, 90t, 104–105. See also Measures; Outcomes; Outputs

Performance partnerships: description and development of, 812; interagency collaborative capacity (ICC), 812–813

PerformWell, 273

Personal Responsibility and Work Opportunity Reconciliation Act (1996), 179b

Pew-MacArthur (2014), 13 Pew Research Institute, 479

Pilot-testing: case studies, 189; focus group interviews, 517–518

Pitfalls (before data collection): 1: failure to assess whether program is evaluable, 703–705; 2: starting data collection too early in the life of the program, 704f, 705; 3: failure to secure input from managers and stakeholders on appropriate evaluation criteria, 704f, 705–706; 4: failure to clarify managers’ expectations about evaluation findings, 704f, 706; 5: failure to pretest data collection instruments appropriately, 704f, 707; 6: use of inadequate indicators of program effects, 704f, 707–708; 7: inadequately training data collectors, 704f, 708

Pitfalls (during data collection): 8: failure to identify and adjust or changes in data collection procedures, 708, 709f, 710f; 9: collecting too many data without adequate time for data analysis, 709f, 710f; 10: inappropriate conceptualization or implementation of the intervention, 709f, 710–711; 11: beginning observation when conditions are at extreme level or not adjusting for this, 709f, 711–712; 12: inappropriate involvement of program providers in data collection, 709f, 712f; 13: overly intrusive data collection procedures that change behaviors of staff or participations, 709f, 712–713; 14: failure to account for drop-of in sample size due to attrition, 709f, 713f; 15: failure to draw representative sample of program participants, 709f, 713f; 16: insufficient number of callbacks to boost response rates, 709f, 714f; 17: failure to account for natural maturation among participants, 709f, 714f; 18: failure to provide a comparison group, 709f, 714–715; 19: failure to take into account contextual factors that affect program outcomes, 709f, 715f; 20: failure to take into account degree of difficulty of helping participants, 709f, 715–716f

Pitfalls (after data collection): 21: overemphasis on statistical significance and under-emphasis on significance of effect size, 716–718; 22: focusing on overall (average) result with inadequate attention to disaggregated results, 717f, 718f; 23: generalizing beyond
confines of the sample or limits of sites included in study, 717, 719–721; 24: failure to acknowledge effects of multiple program components, 717; 721; 25: failure to submit preliminary findings to key staff for reality testing, 717, 721; 26: failure to adequately support conclusions with specific data, 717, 722; 27: poor presentation of evaluation findings, 717, 722–723

Pitfalls in evaluations: after data collection, 716–723; before data collection begins, 703–708; during data collection, 708–716; identifying and managing the most common, 701–703. See also Methodology integrity

Planning. See Evaluation planning

Plausible attribution, 18

Players (stakeholder), 49

PolicyFile, 478

Potential outcomes framework, 144

Power differentials: bases of power-directions of interest diagram to deal with, 48, 51–53; description and potential problems related to, 48; power versus interest grid to deal with, 48–50; stakeholder influence diagram to deal with, 48, 50–51

Power versus interest grid, 48–50

Power writing, 758, 759

Practical program evaluations: culturally responsible, 281–312; description of, 8; improving program performance goal of, 8; increasing demand for information used in, 8; matching evaluation approach to information needs for, 9–17b. See also Program evaluations

Pre-Brown African American evaluators, 306, 310

Pre-post design, 34

Presentations: of evaluation reports, 761–762; formula for successful, 764; Foster Parent Recruitment Report slide presentation, 763e; Golden Rule of giving, 764; tips on preparing overheads and handouts, 762

Pretesting participants, 202–203

Pretesting surveys, 366, 367–368e

Primary intended users: anticipate turnover of, 41; dealing with power differentials among, 48–53; definition of, 37; demonstrating cultural sensitive and competence toward, 40–41; facilitate high-quality interactions with, 40; find tipping point connectors of, 39–40; identifying and working with, 38–41; nurture interest in evaluation, 40; stakeholder identification and analysis of, 41–48. See also Evaluation stakeholders

Privacy issues: agency records and, 336–339; Family Educational Rights and Privacy Act (FERPA), 272, 337; Health Insurance Portability and Accountability Act (HIPAA) on, 211, 212, 272, 337, 819; multisite evaluations (MSEs) and, 247–248. See also Confidentiality issues; Information

Probability sampling, 600

Problem solving: program improvement recommendations for, 730; rapid feedback assessment (TFE) potential problems and, 101–103

Problems: evaluation stakeholder perceptions of, 10; of financing program improvement recommendations, 736, 737f; logic model definition of, 72–74; logic model used to clearly define the, 72–74; potential evaluability assessment (EA), 96–97; power differentials and potential, 48; rapid feedback assessment (TFE), 101–103; role playing, 397–398; summary of agency record potential problems and possible ways to alleviate, 328f; tips on defining the problems addressed by the program, 73e U.S. Department of Education on problems with agency records, 327

Process evaluation studies, 11, 34

Program Assessment Rating Tool (PART), 802

Program data (data study), 190

The Program Evaluation Standards (Yarbrough, Shulha, Hopson, and Caruthers), 291

Program evaluation challenges: controlling the quality of the evaluation process, 816–817; maintaining standards and ethics, 818–820; selecting and training evaluators, 817–818; using evaluation findings to improve programs, 820–822

Program evaluation trends: big data, 825–826; complex adaptive systems, 827; demand for rigorous evidence, 828–830; evaluation mandates, 827–828; information technology (IT), 824–825; visualization, 826

Program evaluations: of community change programs, 259–280; culturally responsible, 281–312; description of, 8; determining the purpose and goals of, 53–55; evaluability assessment (EA) approach to, 88, 90–98; evaluation synthesis approach to, 89, 90, 104; exploratory, 88–106; fieldwork model for, 448–450; government and the politics of, 798–814; how stories enrich, 536b–537; increased demand for rigorous evidence
Program improvement: challenge
Program improvement: evaluation goal, 8
Program structure: logic model on the, 65; resources, activities, and outputs as, 64, 66
Programs: antecedent and mediating factors of, 66; characteristics of, 7; definition of, 54; evaluability assessment (EA) of design of the, 92–93, 95; evaluating community change programs, 239–239; evaluations as not being a substitute for quality implementation of, 830; failure to assess whether it is evaluable, 703–705; selecting for evaluation, 9–10; three approaches to restructuring programs after implementation of, 69–70; see also Managers; Outcomes; Staff; U.S. government programs Promise Neighborhoods program (Department of Education), 259, 266
Propensity score matching (PSM), 151–152
ProQuest Central, 477, 478
PS Power, 169
PsychInfo database, 691
Public Health Service Act, 800
Public Housing program (1937), 260
Public policy research institutes, 479
Purpose network diagram: description of, 46; stakeholder identification and analysis using a, 46–48
Purpose sampling, 186
Pyramid of Strength, 14, 15

Q
Qualitative categories: clustering qualitative codes into, 565–566; definition of, 564
Qualitative codes: categorizing versus, 564; clustering into categories, 565–566; description of, 562, 564–565
Qualitative comparative analysis (QCA), 581–582, 584
Qualitative data: auditability of measurement procedures for, 16–17; clustering codes into categories, 565–566; coding, 562, 564–565; description of, 12, 561; different types of evaluation and uses of quantitative versus, 563–564; sometimes discounted as being valid, 543; value in program evaluation of, 561–562
Qualitative data analysis: descriptive methods for, 566, 567, 572–576; enumerative methods of, 566–572; explanatory method for, 566, 567, 581–584; field studies, 470–471; hermeneutic methods for, 566, 567; 576–581; high quality, 588; multisite evaluations (MSEs), 253–255; overview of methods used for, 566; program

Subject Index

in, 828–830; increasing use of mandates for, 828–828; multisite evaluations (MSEs), 225–256; as not being a substitute for quality implementation of programs, 830; pitfalls in, 701–723; putting it on the Internet, 489–490; rapid feedback evaluation (RFE) approach to, 89, 90, 98–103; recruitment and retention of participants, 197–223; relationship between performance monitoring and, 822–824; small-sample studies approach to, 89, 90, 104–105; theory-driven, 68–70; trends in, 824–830; using stories in, 535–556. See also Evaluation decisions; Practical program evaluations; Timing (evaluation)
Program improvement: challenge of using evaluation findings for, 820–822; making recommendations for, 725–738; as program evaluation goal, 8
Program improvement recommendation strategies: the bottom line on, 738; brainstorm, 731; broader versus narrower view of interpreting the review process, 732; drawing the logic model to reveal the theory of change, 77–81; logic model used to describe theory of change of, 68–70; three approaches to restructuring, 69–70
Programs: antecedent and mediating factors of, 66; characteristics of, 7; definition of, 54; evaluability assessment (EA) of design of the, 92–93, 95; evaluating community change programs, 239–239; evaluations as not being a substitute for quality implementation of, 830; failure to assess whether it is evaluable, 703–705; selecting for evaluation, 9–10; three approaches to restructuring program theories after implementation of, 69–70. See also Managers; Outcomes; Staff; U.S. government programs Promise Neighborhoods program (Department of Education), 259, 266
Propensity score matching (PSM), 151–152
ProQuest Central, 477, 478
PS Power, 169
PsychInfo database, 691
Public Health Service Act, 800
Public Housing program (1937), 260
Public policy research institutes, 479
Purpose network diagram: description of, 46; stakeholder identification and analysis using a, 46–48
Purpose sampling, 186
Pyramid of Strength, 14, 15
Subject Index

evaluation standards and quality criteria for, 589–592; of role playing experiences, 385–386; software used for, 585–587; teams and tasks of the, 587–588. See also Data analysis
Qualitative data collection: community change programs, 271–273; qualitative interviewing, 492
Qualitative themes, 564b
Quality assurance measure: description of, 112, 119; disability adjudicators’ claims processing, 120
Quality of Reporting of Meta-analyses (QUOROM) group, 680
Qualtrics, 369
Quantitative data: community change programs data collection of, 271–273; description of, 12; different types of evaluation and uses of qualitative versus, 563–564; reliability of, 16, 17
Quantitative data analysis: focus group interview used to understand findings of, 508; multisite evaluations (MSEs), 248–253
Quasi-experimental design: community change program evaluation using, 275; definition of, 34
QuestionPro, 369, 487
Questions. See Evaluation questions; Survey questions

R
RAND Corporation, 479, 490
RAND monograph (Schonlau, Ericker, and Elliott), 489
Random sampling, 600
Randomized cohort design, 165
Randomized controlled trial design issues: analytical issues, 171; ascertainment and blinded follow-up, 163; attrition, 164; barriers to the wider use of RCBs, 173–174; biased allocation and secure allocation, 161–162; cluster randomization, 167–169; contamination and cluster randomization, 162–163; crossover and intention to treat (ITT) analysis, 163, 165; example of minimization on two important variables, 168; generalizability or external validity, 172; increased power for very little cost, 170–171; participant preference design, 164–165; quality of randomized trials, 172–173; randomized cohort design, 165; sample size issues, 169–170; waiting list design and stepped wedge design, 165–167
Randomized controlled trials (RCBs): also called an experimental design, 33; description of, 158–159; definitions of, 34, 137; as evaluation gold standard, 13; as “gold standard,” 828–829; history of, 159–160; rationale for randomization, 160; systematic reviews of, 694–695; trial design of, 161–174. See also Comparison groups designs
Rapid feedback assessment (TFE): Comparing evaluability assessment (EA) and, 99; description of, 89; five-step process of, 99–101; issues, problems, and potential solutions when during, 101–103; purpose and cost of, 90; significance of using, 103; tested evaluation designs produced using, 98–99; U.S. Department of Housing and Urban Development (HUD) use of the, 102; when to use, 106
Ratio statistics, 596
Reality testing, 721
Recommendations. See Program improvement recommendations
Recruitment/retention of participants: challenges and potential solutions for, 221–222; cultural considerations of, 219–220; description of, 197–198; for focus group interview, 519–521; frequently asked questions (FAQs) provided to, 214–215; gaining participant cooperation, 212–216; in a health care setting, 210–212; implementing, 209–216; IRB role in protecting rights and welfare during, 203–204, 210, 212, 222, 232, 274, 357, 465; monitoring progress of, 216–219; OBM role in protecting, 203; planning for, 198–203; pros and cons of incentives for, 215–216; reasons for refusing to participate and potential rebuttal approaches, 213; social norms for, 213; staff role in, 204–208
Redundancy (or theoretical saturation), 515
Regression-adjusted covariate design, 146–147
Regression analysis, 609
Regression discontinuity (RD) designs, 34, 153–154
Rehabilitation Act (Section 508), 487
Relevance of Assessment of Culture in Evaluation Institute, 309
Relevance of Assessment of Culture in Evaluation (RACE) [Arizona State University], 285, 309
Reliability: auditability of qualitative data for, 16–17; quantitative data, 16, 17; tips on enhancing, 17; as touchstone of methodological integrity, 703b, 704t, 709t. See also Validity
Subject Index

Remark Web Surveys, 369
Reporting findings. See Evaluation reports
Republic of Georgia, 425
Request for proposals. See RFP (request for proposal)
Request for task order projects (RFTOPs), 773
Research questions. See Evaluation questions
Research Triangle Institute, 490
Resources: description of program, 64, 66; logic model on the, 65/fg for sequential purchase of information process, 89; State Board of Nursing Program Logic on, 114/fg
Respondents. See Survey respondents
RFA (request for application), 774
RFP (request for proposal): description of, 773–774; determining the content of, 774–779; disseminating the, 788–784; forming review panel to conduct independent review of, 784–785; as guideline for selecting the evaluator contractor, 782–789; SOW (statement of work) included in, 776–777; tips on developing a well-defined, 775/fg; writing the, 779–780. See also Evaluation contracting
RFTOPs (Request for task order projects), 773
Risk-adjusted performance measures, 810
River Bluff Orchards, 541
Robert Wood Johnson Foundation (RWJF): Homeless Families Program evaluation by, 227/228; Trust for America’s Health assessment funded by, 96/fg; Urban Health Initiative of, 271, 276b–277b
Role players: attribution problem of, 397; description of, 392–393; determining key characteristics for, 393–394; recruiting and selecting, 394–395; training, 395–396
Role playing: cost considerations, 397; data collection instruments used for, 390–392; ethical and legal issues related to, 406–408; evaluation use of, 385–386; expanding applications for, 405–406; Housing Discrimination Study (2012) using, 389, 393, 395, 399–405; limitations of, 408–409; management and quality control issues of, 396–397; monitoring use of, 386–387; overview of, 383–385; practical problems and solutions, 397–398; practicing enforcement (such as for retail age checks) using, 387; recruiting, selecting, and training role players, 392–396; as a research methodology, 383; statistical analysis of, 398–405
Role playing problems: design deficiencies, 398; detection by test subjects, 397–398; role-player attrition, 397
RwikSurveys, 487

S

Sage Research Methods, 17
Sampling: decisions about trained observer ratings and, 417; evaluation story, 546; failure to account or drop-off in size due to attribution, 709; 713; failure to draw representative of program participants, 709t; 713; multisite evaluations (MSEs) sites for, 234–235; probability, 600; purposive, 186; random, 600; redundancy or theoretical saturation used for focus group interview, 515; role playing, 387–390; square-root law of, 105; stratified, 600; survey, 357–359t; systematic, 600
Sampling distribution, 601
Sampling types: census, 358t; convenience sample, 359t; multistage simple design, 359t; simple random sample, 358t; stratified random sample, 358t; stratified with unequal probability of selection, 358t
San Francisco’s parks inspection program, 425–426
Sar Levitan Center for Social Policy Studies (Johns Hopkins University), 489
SAS software, 617
Satisquest patient survey results summary, 125, 126/fg
Science research methods: qualitative, 12; quantitative, 12; traditional, 11–12
Scotland: Cleanliness Index Monitoring System of, 425; Environmental Protection Act (1990) of, 423; LEAMS (A Local Environmental Audit and Management System) of, 424–425
search.USAGov, 478
Seattle Post-Intelligencer, 539
Section 508 (Rehabilitation Act), 487
Semi-structured interviews (SSIs): analyzing and reporting, 504; description of, 492–493; disadvantages and advantages of, 493–494; drafting questions and the interview guide, 496–500; interview techniques used for, 502–504; selecting respondents and arranging interviews, 495–496; starting the interview, 500–502
Sequential purchase of information process, 89
Service quality measure: description of the, 112, 119; disability adjudicators’ claims processing, 120t
SFI Campbell (formerly the Nordic Campbell Centre), 680
Short-term outcomes: logic model for increasing graduates
Subject Index

Cronbach’s alpha, 17; Sage Research Methods website as resource for, 17; t test, 602, 607–609, 614
Statistics: appropriate reporting of, 617–621; criteria for selecting appropriate, 612–617; descriptive, 597–599; inferential, 599–611; reporting results to high-level public officials, 621–622; Steven’s taxonomy of interval, nominal, ratio, and ordinal, 596–597. See also Causal inferences; Data analysis; Measures; specific taxonomy statistic

STEM program logic model: for increasing middle school STEM program graduates, 75, 76ff–80; performance measurement and evaluation focus on middle school STEM program, 83, 84ff; 85; two levels of detail for increasing graduates entering STEM careers, 79ff
Stepped wedge design, 165, 166–167
Stories. See Evaluation stories
Story examples: ECHO, 541; “medical mistakes,” 538ff; “understanding others,” 536ff
Strategy maps, 115
Stratified sampling, 600
Subjects (stakeholder), 49
Substance Abuse Mental Health Services Administration (SAMHSA) Homeless Families Program, 229
Summative evaluation, 11ff
SumQuest, 369
Survey Gizmo, 369, 487
Survey modes: computer-assisted telephone interviewing (CATI or CAPI), 370, 371, 373, 374, 376; employ quality control with each, 372–375; face-to-face, 351, 353, 370, 371, 374; mail, 351, 352, 366, 368; mixed-mode, 353–354;

860

entering STEM careers, 76ff; State Board of Nursing Program, 114ff
Significance (statistical significance), 10, 611
Single-case case study design, 185
Slope coefficients, 610
Small-sample studies: description off, 89; purpose and cost of, 90; vetting performance measures using, 104–105; when to use, 106ff
SmartDraw, 81f
Snap Survey Software, 360
Social Security Act Title V, 800
Social Security Amendments (1967), 800
Social Solution’s Efforts to Outcomes, Salesforce, and Community Techknowledge (or CTK), 273
Society for Prevention Research, 679
SOW (statement of work), 776–777
Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), 807, 808
Specific stakeholders, 37
Sponsors. See Evaluation sponsors
SPSS (Statistical Package for the Social Sciences), 607, 613, 616, 617
Square-root law, 105
Staff: case study training and qualifications of, 188; communication skills of, 206–207; consideration of personal and professional background of, 205–206; field visit assignments given to, 464–465; guide for evaluability assessment (EA) review of interviews with, 94ff; interpersonal qualities of, 206; overly intrusive data collection procedures that change behaviors of, 709ff; 712–713; participant recruitment and retention of, 204–208; performance incentives for government program, 810; training and supervision of, 207–208; training case study, 188. See also Managers; Programs
Stakeholder identification and analysis: choosing participants for, 44–46; conducting basic, 43–44; creating a purpose network diagram for, 46–48; techniques for, 41–43
Stakeholder influence diagrams, 50–51
Stakeholders. See Evaluation stakeholders
Star diagrams, 51
STATA software, 617
State Board of Nursing Program Logic, 114ff
Statistical conclusion validity, 20, 703ff, 704ff, 709ff
Statistical confidence level, 604–605
Statistical hypothesis testing: chi-square test for, 602, 605–607, 614ff; design features likely to generate false positives or negatives, 603ft; type I error (false positive) and type II error (false negative), 602–603ft; multiple regression model for, 609–611; overview of, 601–603; practical significance, 611; regression analysis for, 609; selecting a statistical confidence level for, 604–605; the t test for, 602, 607–609, 614ff; using a confidence interval to convey results, 605
Statistical Package for the Social Sciences (SPSS), 607, 613, 616, 617
Statistical significance, 10, 611
Statistical techniques: matching analytical objectives to, 614ff; other factors affecting selection of, 615–617; selecting to sort measures or units, 613, 615
Statistical tests: chi-square test, 602, 605–607, 614ff;
Subject Index

other methods or, 354; telephone, 351t, 352, 370–371, 374; web surveys, 351t, 355, 369–370, 479–489
Survey Monkey, 369, 487
Survey Nonresponse (Groves, Dillman, Eltinge, and Little), 375
Survey planning: categories and criteria for a survey design matrix, 347e; decide on analysis plan, 350; decide on plan for collecting the data, 350–354; determine need and feasibility of survey, 346, 349; determine population of interest, 349–350; establish evaluation questions, 346; example of a survey design matrix, 348e–349e; identifying who will conduct the survey, 355; selecting the right consultant or contractor, 356e; when to contract out, 355e
Survey questions: common mistakes in writing, 365e; considerations for the “no opinion” or “don’t know” option, 365e; crafting good, 361–363, 365; pretesting, 366, 367e–368e; tips on sequencing, 364e; wordings to measure attitudes and perceptions reliably, 363e–364e; writing web survey, 482–483e. See also Evaluation questions
Survey reports: contingency table consolidating multiple variables, 380e; contingency table presenting survey finding, 380e; presenting, 378–379, 381; tips on designing effective tables, 378e–379h
Survey respondents: cognitive interviewing of, 367e–368e; collecting data from, 366, 368–371; considering the target population, 360; contacting web survey, 487–489; craft good questions for, 361–365e; drafting questions and the interview guide for interviewing, 496–500; employ quality control when gathering data from, 372–375; get a foot in the door, 361; insufficient number of callbacks to boost response rates of, 709; 714; response rates, 375–376; selecting semi-structured interviews, 493–496; starting the interview with, 500–502; tips on incentives, 361e; validating web survey representation by, 484. See also Respondents; Survey respondents
SurveyConsole, 487
SurveyMethods, 487
Surveys: collect data from respondents, 366–376; community change programs use of, 272; design the survey instrument, 359–366; overview of, 344–345; planning the, 346–357; prepare data for analysis, 376–377; present findings, 378e–381; pretesting, 366, 367e–368e; select the sample, 357–359e; training interviewers, 371–372
SurveySaid, 369
SurveyShack, 487
Survs, 487
Swiss Environmental Impact Assessment (EIA), 582, 584
Synthetic reviews, 676
Systematic reviews: academic disciplines and policy sector dependence on, 693–694; by-products of, 694–695; definition of, 678; examples off inclusion criteria for, 683–684; interpreting and reporting, 688–689; of multiple evaluations, 674–675; of randomized controlled trials (RCTs), 694–695; resources employed to do a good job on, 689–692; role of bias in, 673–674; search strategies used for, 681–683; steps for producing a, 680–689; value added and usefulness of, 692–693
Systematic sampling, 600
Task Force on Community Preventive Services (Community Guide initiative), 690
Teach for America (TFA), 150
Teachers: Cincinnati Public Schools Teacher Evaluation System (TES) program for observing, 429–430; 431e–432e; scoring guide using four-point scale for evaluating, 431e–432e; trained observer ratings of, 428–433
Technical report: description of the, 746, 752; tips on producing a, 753e
Technology: complex adaptive systems, 827; evaluation trend toward increasing use of IT, 824–825; GIS and data visualization, 826; overlap of evaluation contracting services with IT services, 768; software applications for web survey, 369, 487; software packages for data analysis, 246, 254, 607; software packages for qualitative data analysis, 583–587; Statistical Package for the Social Sciences (SPSS), 607, 615, 616, 617; for trained observer ratings, 441. See also The Internet
Telephone communication, 210
Telephone surveys: collecting data from respondents using, 370–371; compared to other survey modes, 351t, 352;
Subject Index

computer-assisted telephone interviewing (CATI), 370, 371, 373, 374, 376; sample contact record sheet for, 374

telling training’s story (Brinkerhoff), 540

temporary assistance for needy families (TANF), 189, 272

ten-page reports, 746, 750, 752

Tennessee Class Size randomized trial, 692

text enhancements, 753

TFAH (The Trust for America’s Health), 96

themes (qualitative), 564

theoretical saturation (or redundancy), 515

theory-based evaluation (TBE), 34

theory-driven evaluation, 68–70

theory of change: comprehensive community initiatives (CCIs) application of, 263; drawing the logic model to reveal the program, 77–81; logic model used to describe the, 68–70

thought leaders, 744

3ie, 681

timing (evaluation): contracting concept plan on products and schedule and period of performance, 771–772; determining for community change programs, 259–271; of either CBA or CEA, 639–640; field protocol for on-site schedule, 467; importance of, 10; of survey data collection, 356–357; when evaluator should make recommendations based on evaluation, 725–726. See also Program evaluations

tipping point connections, 39–40

the tipping point (Gladwell), 535

title V (social security act), 800

trained observer contractors: New Yorkers for Parks (NY4P), 427–428; Worcester Regional Research Bureau (WRRB), 426–427

trained observer initiatives: decisions about ratings and sampling, 417; evaluating judges in their courtrooms, 433, 434; evaluating teachers in their classrooms, 428–433; key requirements of effective, 440; key requirements of effective initiatives for, 440; presenting findings of, 435–438; step-by-step operation of, 433; what you will need to start, 416

trained observer programs: Cincinnati Public Schools Teacher Evaluation System (TES), 429–430; City of Durham, 422–423; ComNET2Go initiative, 421–422; Computerized Neighborhood Environment Tracking (ComNET), 418–421; description of, 417; Fund for Modern Courts, 433, 434; Keep America Beautiful (KAB), 417–418; LEAMS (A Local Environmental Audit and Management System) [Scotland], 423–425; New York City’s Mayor’s Office of Operations, 423; Objective Structured Clinical Examination (OSCE) for medical students’ evaluations, 433, 435; outsider assessors running, 426–428; San Francisco parks inspection program, 425–426

trained observer ratings: benefits of using, 441–442; decisions about sampling and, 417; examples of rating scales items, 436–437; intensity map of, 439/fig; presenting findings, 435–438; quality control, 438–440; questions to address before adopting and designing, 414; sample report from, 438; technology versus paper used for, 441; uses for, 413; what you will need to start, 416; when it is appropriate to use, 414–416

Training: challenge of evaluator selection and, 817–818; field visit, 466; inadequate data collector, 704; multisite evaluations (MSEs), 243–244; role players, 395–396; staff supervision and, 207–208; survey interviewers, 371–372

transferability of findings, 19

Treasury’s New Market Tax Credit program, 262

triangulation, 191

tweets in bahasa, 569, 570

two-page executive summary: description of, 746, 749–750; Foster Parent Recruitment Report example of, 754–755; tips on writing the, 751

2ask.net, 487

type I error (false positive), 602–603

type II error (false negative), 602–603

U

unicef, 569

unit of analysis: assessing stability or mobility of the, 268–269; determining case study design, 184–185; field study site selection and, 456–457; selected for evaluating community change programs, 265–266. See also Data analysis

united nations global pulse lab, 569

united way of america, 67, 78

univariate statistics, 597

university of minnesota extension service, 540

urban health initiative (RWI), 271, 276b–277b

Urban Institute (UI), 96, 425, 460, 479, 490

urban renewal program, 260
Subject Index

863

Urban school talent development intervention program evaluation, 299–301
U.S. Bureau of Community Health Services, 807
U.S. Census Bureau, 357
U.S. Code, 479
U.S. Congress, 828
U.S. Department of Agriculture, 807
U.S. Department of Education: Family Educational Rights and Privacy Act (FERPA) enforced by the, 272, 337; Promise Neighborhoods program of the, 259, 266; regarding problems with agency records as data sources, 327; reporting on 2012 rates used as “status dropout rate,” 333; What Works Clearinghouse (WWC) of the, 679, 680, 682, 683, 684, 686, 690, 691
U.S. Department of Health and Human Services: Code of Federal Regulations by the, the, 332; Foster Care Parent Recruitment Report by the, 747, 749b, 750b–751b, 754a–755a, 757–758fg, 763g; Healthy People program of, 813; increased use of evaluation by the, 800; “promote safety” goal of, 113, 115; role-playing program established by the, 405
U.S. Department of Health, Education, and Welfare (now the Department of Health and Human Services), 800
U.S. Department of Housing and Urban Development (HUD): Choice Neighborhoods program of the, 259; Community Development Block Grant of the, 276; Jobs Plus initiative of, 275; rapid feedback evaluation of a housing production program of the, 102e; role-playing study to estimate housing discrimination against the deaf, 391
U.S. Department of Labor, 800, 807
U.S. Department of Veterans Affairs, 540
U.S. Employment Service, 800
U.S. Environmental Protection Agency, 149
U.S. Government Accountability Office (GAO): “Clean Air Act: Observations on EPA’s Cost-Benefit Analysis of its Mercury Control Options” by, 670c; compliance reviews governing evaluator recommendations at, 726; evaluation design process used by, 22b–25b; evaluation synthesis used by the, 104; on exploratory case studies, 178; external reviews of agency’s evaluation activities suggested by, 817; GAO Report as “gray literature,” 478; on getting the best data analysis, 192; on purposive sampling for case studies, 186; sample design matrix used by, 26fg, 182; selection of evaluation type by, 12; systematic review of studies of interventions on WIC dependency by the, 684; on taking into account contextual factors that affect program outcomes, 709, 715; on tips for avoiding common case study pitfalls, 194; on triangulation used for data analysis, 191; Yellow Book standards of the, 820
U.S. Government Printing Office, 478
U.S. government programs: creating incentives for high performance of, 809–811; creating incentives for high program performance, 809–811; developing agreement on key performance indicators, 811–812; developing performance partnerships, 812–813; evaluation used to increase transparency, accountability, and improve performance, 799; federal performance portal of, 803; overcoming political and bureaucratic challenges of evaluation use by the, 805–813; political and bureaucratic challenges affecting use of evaluation by the, 804–805; redesigning agency management systems to focus on results, 806–809. See also Agency records; Programs
U.S. legislation: Affordable Care Act, 405; Americans with Disabilities Act, 391; Chief Financial Officers Act (1990), 801; Child Health Act (1967), 800; Economic Opportunity Act (1964), 799, 800; Elementary and Secondary Education Act (1965), 799–800; Fair Housing Act, 389, 391; Family Educational Rights and Privacy Act (FERPA), 272, 337; Government Performance and Results Act (GPRA), 770, 801, 821; GPRA Modernization Act (2010), 802–804, 821; Health Insurance Portability and Accountability Act (HIPAA) on, 211, 212, 272, 337, 819; Home Mortgage Disclosure Act, 272–273; Housing Act (1949), 260; No Child Left Behind Act (2002), 118, 298, 802; Paperwork Reduction Act (PRA), 203; Personal Responsibility and Work Opportunity Reconciliation Act (1996), 179c; Public Health Service Act, 800; Rehabilitation Act (Section 508), 487; Social Security Act Title V, 800; Social Security Amendments (1967), 800. See also Legal issues
Subject Index

V

Validity: authenticity of measurement or, 15; culturally responsible evaluation (CRE) justifications of, 304t–305t; data analysis issue, 15–16, 18–20; external or generalizability, 19, 20t, 172, 702t; face, 15; internal (confirmability), 18–19t, 65f, 161–172, 702t; measurement, 702t, 704t, 709t; statistical conclusion, 20, 703t, 704t, 709t; testing measures for, 15–16. See also Reliability

Value-added designs: adjusted for additional covariates, 147–148; regression adjusted for a preprogram measure, 145–146

Vanderbilt University’s Biostatistics Department, 169

Veterans Health Administration (VHA), 808–811

Veterans Integrated Service Networks (VISNs), 808

Virginia Performs, 811

Visualization using GIS platforms, 826

W

Waiting list design, 165–167

The Wall Street Journal, 693

Washington Families Fund High Needs Family evaluation, 227t, 228

Washington Institute of Public Policy, 928

Washington State Institute for Public Policy (WSIPP), 690, 694

Web surveys: compared to other modes of surveys, 351t, 353, 369–370; contacting respondents, 487–489; drafting questions for, 482–484; outsourcing, 486–487; pros and cons of conducting, 479–482; using unique design aspects of, 484–486; validating respondent representation, 484

WebSurvey Methodology website, 369

Westinghouse study of Head Start, 151

WFF High Need Family MSE, 227t, 232, 234, 253

What Works Clearinghouse (WWC), 679, 680, 682, 683, 684, 686, 690, 691

White House Office of Management and Budget (OMB), 203

WIC (Special Supplemental Nutrition Program for Women, Infants, and Children), 807, 808

William and Flora Hewlett Foundation’s Neighborhood Improvement Initiative, 277t–278t

W.K. Kellogg Foundation, 64

Worcester Regional Research Bureau (WRRB), 426–427

Work Incentive Program (WIN) field studies, 447t, 450, 460–461

WorldCat, 477–478

Writing style and layout: color used in, 758–759; electronic reports, 760–761; graphs, tables, and other large graphics, 756, 757; making presentations, 761–764; overview of the, 752; paper copies and tips on helping others to make copies of your report, 760t; physical considerations of, 758; power writing, 758, 759t; types on text enhancements, 753t

Y

Yahoo! search engine, 476

Yellow Book standards (GAO), 820

YMCA programming: contingency table: participants’ preferences by gender of child, 606t; observed programming preferences by gender of child, 606–607t; reported number of years participating by child gender and favorite, 608t

Z

Z model: for increasing number of graduates entering STEM careers, 79f; relationship between performance measurement and, 83

Zoomerang, 369