INTRODUCTION
Mixed methods is a methodology that focuses on collecting, analyzing, and mixing both quantitative and qualitative data in a single study or multiphased study. Its central premise is that the use of quantitative and qualitative approaches in combination provides a better understanding of research problems than either approach alone. Although there is some debate as to whether one approach must, of necessity, assume the dominant role and the other approach a supplementary role in a research project, there is general agreement that the critical feature of mixed methods is the integration of quantitative and qualitative methods; conducting one or more qualitative studies and one or more quantitative studies in the same project without integrating the methods or the results (i.e., engaging in “parallel play”) is not mixed methods per se, but is perhaps better described as a multimethod design. In a mixed method design, each set of methods plays an important role in achieving the overall goals of the project and is enhanced in value and outcome by its ability to offset the weaknesses inherent in the other set and by its “engagement” with the other set of methods. Consequently, mixed methods represent both a model of and a model for transdisciplinary research. In the past decade, mixed methods have come to play a critical role in the field of dissemination and implementation (D&I) science. A recent study found that 69% of 67 D&I research grants funded by the National Cancer Institute between fiscal years 2010 and 2012 planned to use both qualitative and quantitative methods. This role has emerged from both necessity and opportunity. Similar to the use of hybrid designs in which evidence-based practice effectiveness and implementation are addressed simultaneously, mixed methods are often used to simultaneously answer confirmatory and exploratory research questions, and therefore verify and generate research in the same study. As D&I science is a relatively “new” discipline, generating theory has been accorded the highest priority. Some of the theories, frameworks, and models that have been developed explicitly call for the use of both quantitative and qualitative methods due to the complexity of the subject matter, the importance of understanding both general principles and specific context, and the need to acquire depth as well as breadth of understanding of D&I. Further, “empirical research is needed to study how and the extent to which implementation theories, models and frameworks contribute to more effective implementation and under what contextual conditions or circumstances they apply (and do not apply). It is also important to explore how the current theoretical approaches can be further developed to better address implementation challenges. Hence, both inductive construction of theory and deductive application of theory are needed.”

In D&I science, mixed methods are most commonly used to identify barriers and facilitators to successful implementation, but may also be used as a tool for developing strategies and conceptual models of implementation and sustainment, monitoring the implementation process, and enhancing the likelihood of successful implementation and sustainment. Qualitative methods are generally used inductively to examine the context and process of implementation with depth of understanding, while quantitative methods are commonly used deductively to examine the content and outcomes of implementation with breadth of understanding. Following
the examples of nursing, evaluation, public health, health services, education, and the social and behavioral sciences, there are several typologies and guidelines for the use of mixed methods designs in D&I science.

The aims of this chapter are to: (1) provide a brief overview as to the structure, function, and process of mixed methods; (2) describe what quantitative and qualitative methods can and cannot do within the context of D&I science; and (3) provide examples of what mixed methods can accomplish in implementation science.

**Characteristics of Mixed Methods Designs**

**Structure**

Mixed method designs in D&I research can be categorized in terms of their structure, function, and process. Quantitative and qualitative methods may be used simultaneously or sequentially, with one method viewed as dominant or primary and the other as secondary, although equal weight can be given to both methods. A review of published studies of D&I found that most studies involved the simultaneous use of quantitative and qualitative methods, and most used quantitative methods as the primary or dominant method and qualitative methods as the secondary or subordinate method. However, a little less than half of the studies used balanced designs in which quantitative and qualitative methods were used simultaneously and given equal weight. For instance, Whitley and colleagues documented the process of implementation of an illness management and recovery program for people with severe mental illness in community mental health settings using qualitative data to assess perceived barriers and facilitators of implementation and quantitative data to assess implementation performance based on assessments of fidelity to the practice model, with no overriding priority assigned to either aim. Some studies gave equal weight to qualitative and quantitative data for the purpose of evaluating fidelity and implementation barriers/facilitators even though the collection of qualitative data to assess implementation was viewed as secondary to the overall goal of evaluating the effectiveness of an intervention.

**Function**

In D&I research, mixed methods have been used to achieve one or more of five different functions (Table 20.1). First, qualitative and quantitative methods may be used sequentially or simultaneously to answer the same question. This is known as convergence. There are two specific forms of convergence: triangulation and transformation. Triangulation involves the use of one type of data to validate or confirm conclusions reached from analysis of the other type of data. Swain and colleagues used triangulation to identify commonalities and disparities between quantitative data obtained from closed-ended questions and qualitative data obtained from open-ended questions of a survey administered to 49 participants, each participant representing a distinct practice site. Aarons and colleagues compared findings obtained from analyses of quantitative and qualitative data to assess convergence of assessments of feasibility, acceptability, and utility of a leadership implementation intervention. Transformation involves the sequential quantification of qualitative data or the use of qualitative techniques to transform quantitative data. The technique of concept mapping, used by Aarons and colleagues, where qualitative data elicited from focus groups are “quantitized” using multidimensional scaling and hierarchical cluster analysis to identify stakeholder perceptions of barriers and facilitators of implementing evidence-based practices in community mental health settings, is an example of transformation. In another illustration, Watts and colleagues conducted semistructured interviews with staff at participating clinics using the Promoting Action on Research Implementation in Health Services (PARIHS) framework to develop overarching questions in a study of the implementation of evidence-based psychotherapies for PTSD in VA specialty clinics. Transcripts of these interviews were coded by domain and element of the PARIHS framework, and a scoring rubric was used to transform each element of the framework into a numeric value. Damschroder and Lowery embedded the constructs of the Consolidated Framework for Implementation Research in semistructured interviews and then assigned numerical ratings that reflected their valence (positive or negative influence) and their magnitude or strength. Qualitative comparative analysis is another form of mixed method that “can be used to (1) analyze small to medium number of cases (i.e., 10 to 50) when traditional statistical methods are not possible, (2) examine complex combinations of explanatory factors associated with translation or implementation “success,” and (3) combine quantitative and qualitative data using a unified and systemic analytical approach.”

5 (p.201)
Second, quantitative and qualitative methods may be used to answer related questions for the purpose of evaluation or elaboration. This is known as complementarity. In evaluative designs, quantitative data are used to evaluate outcomes while qualitative data are used to evaluate process. In elaborative designs, qualitative methods are used to provide depth of understanding to complement breadth of understanding afforded by quantitative methods. Qualitative methods used to study process and context and quantitative methods to study outcomes.

Gilburt et al. used mixed methods to achieve complementarity in their evaluation of implementation of a recovery-oriented practice through training across a system of mental health services, using a quantitative audit of care plans in a random sample of 700 patients to assess change in core plan topics and in responsibility of action and semi-structured interviews with team leaders to explore understanding of recovery, implementation within the service and the wider system, and perceived impact of the training on individual practice and that of the team.

Kramer and Burns used data from qualitative interviews with providers as part of a summative evaluation to understand the factors contributing to partial or full implementation of a CBT for depressed adolescents in two publicly funded mental health care settings. Brunette and colleagues used qualitative data collected from interviews and ethnographic observations to elucidate barriers and facilitators to implementation of integrated dual disorders treatment and explain differences in treatment fidelity across the study sites. Duffy and colleagues used the qualitative data obtained from a smaller number of nurse interviews to further explain the more close-ended, but more generalizable, patient and nurse surveys and EMR data that were obtained from larger numbers of participants.

<table>
<thead>
<tr>
<th>TABLE 20.1 FUNCTIONS OF MIXED METHOD DESIGNS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design type</strong></td>
</tr>
<tr>
<td>Convergence</td>
</tr>
<tr>
<td>Complementarity</td>
</tr>
<tr>
<td>Expansion</td>
</tr>
<tr>
<td>Exploratory/Development</td>
</tr>
<tr>
<td>Sampling</td>
</tr>
</tbody>
</table>
Fourth, one set of methods may be used to answer questions that will enable use of the other method to answer other questions. This is known as development. In implementation research, there are three distinct forms of development: instrument development, conceptual development, and intervention development or adaptation. Zazelli and colleagues connected qualitative data collected from semistructured interviews with 15 program administrators to the development of a conceptual model of implementation of Functional Family Therapy that could then be tested using quantitative methods. Blasinsky et al. used qualitative data obtained from semistructured interviews to develop a rating scale to construct predictors of program outcomes and sustainability of a collaborative care intervention to assist older adults suffering from major depression or dysthymia. Sales and colleagues describe a protocol for using quantitative outcomes data and qualitative data on implementation barriers and facilitators to design feedback reports for staff as a strategy for implementing goals of care conversations with veterans in VA long-term care settings.

Finally, there is the sequential use of one method to identify a sample of participants for use of the other method. This is known as sampling. Aarons and Palinkas purposively sampled 15 case managers selected to represent those having the most positive and those having the most negative views of the Safe Care® home-based intervention based on results of a web-based quantitative survey asking about the perceived value and usefulness of Safe Care®. Woltmann et al. used qualitative data obtained through interviews with staff, clinic directors, and consultant trainers to create categories of staff turnover and designations of positive, negative, and mixed influence of turnover on implementation outcomes. These categories were then quantitatively compared with implementation outcomes via simple tabulations of fidelity and penetration means for each category.

**Process**

The process of integrating quantitative and qualitative data occurs in three forms, merging the data, connecting the data, and embedding the data (Figure 20.1). In implementation research, merging the data occurs when qualitative and quantitative data are brought together in the analysis phase to answer the same question through triangulation or related questions through complementarity. Connecting the data occurs when the analysis of one dataset leads to (and thereby connects to) the need for the other dataset, such as when qualitative results lead to the subsequent collection and analysis of qualitative data (i.e., expansion) or when qualitative results are used to build to the subsequent collection and analysis of quantitative data (e.g., development). Embedding the data occurs when qualitative or mixed method studies of treatment or implementation process or context are embedded within larger quantitative studies of treatment or implementation outcome for the purpose of complementarity, convergence, or expansion. In general, quantitative and qualitative data are merged when the two sets of data

![Three ways of mixing quantitative and qualitative data](image_url)
are used to provide answers to the same questions, connected when they are used to provided answers to related questions sequentially, and embedded when they are used to provide answers to related questions simultaneously. In describing a protocol for an implementation science study of HIV-related stigma in communities and health settings within a large, pragmatic cluster-randomized trial of a universal testing and treatment intervention for HIV prevention in Zambia and South Africa, Hargreaves and colleagues propose to collect quantitative data on indicators of HIV-related stigma from large probability samples of community members, health workers, and people living with HIV, and qualitative data including in-depth interviews and observations from members of these same groups. These data will be used to accomplish three specific aims: (1) compare HIV-related stigma measures between study arms; (2) link data on stigma to measures of intervention implementation; and (3) explore changes in the dominant drivers and manifestations of stigma in study communities and the health system.

**Operation**

An illustration of how and why quantitative and qualitative methods are mixed can be found in a randomized cluster trial study protocol of an intervention model, Implementing School Nursing Strategies to Reduce LGBTQ Adolescent Suicide. Mixed methods will be used to examine individual, school, and community factors influencing both implementation process and youth outcomes. A cluster randomized controlled trial will quantitatively assess whether LGBTQ students and their peers in Reduce LGBTQ Adolescent Suicide (RLAS) intervention schools report reductions in suicidality, depression, substance use, bullying, and truancy related to safety concerns compared with those in usual care schools. Implementation progress and fidelity for each evidence-based (EB) strategy in RLAS intervention schools will be measured using a modified version of the Stages of Implementation Completion checklist. During the implementation and sustainment phases, annual focus groups will be conducted with the RLAS intervention schools to document their experiences identifying and advancing adaptation supports to facilitate use of EB strategies and their perceptions of the Dynamic Adaptation Process.

"Data will be merged by the following: (a) linking qualitative and quantitative databases; and (b) embedding one within the other so that each play a supporting role for the other. Results of each data set will be placed side-by-side to examine: (1) convergence—do interview data concur with New Mexico Youth Risk and Resilience Survey data regarding the impact of the intervention on suicide risk among LGBTQ students?; (2) expansion—can web-based survey data that suggest a disempowering school climate be explained by qualitative interview data?; and (3) complementarity—do the qualitative results describe contextual factors that reflect the variability represented by confidence intervals or variance estimates in statistical analyses of "bullying on school grounds" and "fear-based bullying"?

The protocol for examining individual and organizational factors that affect implementation of interventions for children with autism in public schools described by Locke and colleagues is also explicit in the structure, function, and process of the mixed-method design. Researchers will employ a sequential (QUAN → QUAL) design, giving equal weight to the two sets of methods, beginning with a quantitative assessment of individual (e.g., attitudes of EBIs) and organizational (e.g., culture and climate, leadership) characteristics as predictors of implementation and sustainment of autism interventions, followed by qualitative methods to understand the appropriateness and fit of four autism EBIs within the school context, and principals’ and teachers’ experiences and perspectives regarding the implementation process in a subset of classrooms that are either high or low performing based on their fidelity data. The function of the design is evaluative and elaborative complementarity (quantitative data to assess outcomes and qualitative data to assess implementation process), and the process is connecting the two data sets so that the qualitative data builds upon the findings of the quantitative data.

**WHAT QUANTITATIVE METHODS CAN AND CANNOT DO IN IMPLEMENTATION SCIENCE**

"Quantitative research is a mode of inquiry used often for deductive research, when the goal is to test theories or hypotheses, gather descriptive
Quantitative methods have specific characteristics that affect the role they play in D&I research. As the name implies, quantitative methods focus on the quantity of phenomenon being studied, mostly through the analysis of numeric data to determine statistical significance of the associations between constructs or the differences across groups. \(^{20,21}\) Given the recent explosion of conceptual models and frameworks proposed to explain the characteristics of and processes by which successful D&I occur,\(^{13,54}\) quantitative methods are often used deductively to test assertions and confirm hypotheses based on existing conceptual frameworks and models.

Many of these frameworks and models describe implementation in terms of key stages or processes.\(^{55}\) One reason to use quantitative methods is to quantify or characterize these stages and processes. For example, recent studies have used finite mixture modeling to identify and characterize subgroups with distinct patterns of intervention participant attendance across time\(^{56}\) and implementation quality.\(^{57,58}\) McIntosh and colleagues\(^{56}\) found four distinct patterns of implementation in a large sample of schools implementing schoolwide positive behavioral support framework across 5 years: Sustainers (schools likely to meet fidelity criterion all 5 years), Slow Starters (schools with inconsistent fidelity across the first 3 years, but higher likelihood of meeting fidelity criterion in years 4 and 5), Late Abandoners (schools likely to meet fidelity criterion in first 3 years and unlikely in years 4 and 5), and Rapid Abandoners (schools likely to meet fidelity criterion during the first year, followed by rapid decline across the remaining 4 years). Quantitative methods were also used in these studies to examine predictors of the identified patterns\(^{56,58}\) and to determine how the patterns related to other implementation and intervention outcomes.\(^{57}\)

Implementation frameworks and models also posit the community, organizational, and intervention factors associated with successful implementation\(^{11}\) and/or the links between aspects of implementation and participant outcomes.\(^{59}\) Therefore, quantitative methods are also used to test hypotheses related to predictors of successful implementation\(^{58,60–63}\) and to quantify implementation barriers and facilitators.\(^{64}\) In a cross-sectional study, Shapiro et al\(^{64}\) surveyed 174 mental health service providers who were trained within the last 2 years to implement 'Triple P,' an evidence-based parenting intervention. They used logistical and linear regression analysis to determine whether provider attributes, perceived barriers and facilitators to program use, and attitudes toward use of evidence-based practice predicted providers’ self-reported use of the intervention following initial training. Aarons and colleagues\(^{65}\) also used quantitative methods to survey service providers about the factors predictive of continued use of a parenting intervention. The authors used ordinal regression analyses to determine how different types of administration leadership (e.g., transformational, transactional, passive-avoidant) within the inner and outer context of the intervention were associated with the organizations’ level of program sustainment (full, partial, or nonsustainment).

The theoretically based assertions about how different aspects of implementation lead to intervention outcomes can also be tested using quantitative methods.\(^{57,58,60–66}\) Goncy et al\(^{66}\) examined the relative impact of teachers’ instructional adherence, procedural adherence, and quality of delivery of a classroom-based bullying prevention program on students’ responsiveness to the intervention. These variables were quantitatively assessed through 288 observations of 44 teachers’ implementation of the program over 2 years. The authors used multilevel models, which accounted for clustering of multiple observations for the same teacher across time and controlled for important covariates (e.g., school, intervention year, number of years teacher had experience with intervention), to determine whether the different aspects of adherence and quality of delivery were significantly associated with student responsiveness above and beyond other confounds. In a study by Berkel et al.,\(^{45}\) several fidelity scores were
devised to determine how fidelity-specificity was related to participant outcomes in an evidence-based family intervention for rural African American youth and their caregivers. Fidelity was operationalized as total number of fidelity items completed, divided by total possible and multiplied by 100 to produce a percentage of content delivered. The authors conducted a series of three regression analyses, with fidelity scores predicting change in youth racial socialization—each analysis was designed to examine fidelity’s impact at increasing degrees of specificity for the targeted outcome: average fidelity across all sessions, average fidelity for the session on racial socialization, average fidelity for the activity in that session on racial socialization.

Another goal of D&I science is to develop and evaluate discrete strategies aimed at effectively moving evidence-based health interventions into routine use in routine care settings. Quantitative methods, especially within the context of randomized controlled designs, are used to determine the impact implementation strategies have on measureable implementation (e.g., adoption, fidelity, sustainability) and intervention outcomes (e.g., participant health outcomes). For example, Atkins and colleagues designed a study to test the impact of using peer-identified teachers (i.e., key opinion leaders) for disseminating information on educational and mental health best practices for students with attention-deficit/hyper-activity disorder (ADHD). Schools were randomly assigned to one of two conditions: key opinion leader plus mental health consultants or mental health consultants only. Mixed effect regression models were used to determine whether teacher-reported use of the ADHD best practices varied by condition, and whether these differences could be explained by the support teachers received from the key opinion leader in their school. In another randomized controlled trial, Lochman and colleagues designed a study to test the impact of two different training models (high intensity vs. normal intensity) for school counselors implementing Coping Power, an evidence-based aggression prevention program. Fifty-seven school counselors were randomly assigned to one of three conditions: high intensity (basic training plus ongoing, individualized monitoring, technical assistance and feedback), normal intensity (basic training only), or a comparison condition (no training). They used hierarchical linear modeling to compare the changes in student behavioral outcomes in schools where counselors received high, normal, or no training, to determine whether training intensity made a difference for intervention outcomes.

Glisson et al. also conducted a randomized controlled trial to test the added benefit of combining an intervention strategy with an existing evidence-based treatment (Multisystemic Therapy, MST). In this case, the implementation strategy being tested was ARC (availability, responsiveness, and continuity), an organizational intervention designed to improve community-based mental health service delivery by developing social networks with key community stakeholders, and identifying and removing service barriers. They used a 2 × 2 factorial design to first randomly assign counties to ARC, the organizational intervention, and then randomly assigned delinquent youth within those counties to receive MST, the evidence-based treatment, or not. This resulted in four conditions: MST plus ARC, MST only, ARC only, or control. The researchers used multilevel, mixed effect regression analyses to compare youth behavioral and service outcomes (i.e., out-of-home placements) across all four conditions. They also examined if MST fidelity levels significantly varied across the two conditions implementing MST in an attempt to determine if the effects of ARC could be explained by improvements in MST fidelity.

Another reason quantitative methods are used in D&I science is for summarizing quantitative findings across multiple studies in systematic reviews and meta-analyses. These reviews are often used to determine the degree to which specific types of implementation strategies or interventions are effective across studies and settings. For example, Powell and colleagues conducted a systematic review of 11 studies testing the effectiveness of implementation strategies for facilitating the delivery of evidence-based mental health treatments. They used vote counting to determine the proportion of studies that found statistically significant positive results of the strategy on implementation or clinical outcomes. Meta-analysis uses more formalized quantitative techniques to summarize quantitative findings across studies. For example, Webb and colleagues identified 85 randomized controlled trial studies (N = 42,236 participants) testing the effectiveness of Internet-based health behavior change interventions. They used Hedges g (the difference between treatment and control means, divided by the pooled standard deviation) to determine the effect size for each intervention.
and then computed a weighted average effect size across all interventions on participants’ health behavior. Systematic reviews and meta-analyses are also helpful in determining how similar or different results are across studies and synthesizing findings related to specific implementation frameworks and models.65,78

As illustrated in the studies described, quantitative methods play an important role in D&I science. However, if used in isolation, they are likely to hinder the field’s advancement due to several important limitations. Some of these limitations are driven by the fact that quantitative methods are only as good as the research designs, measures, and statistical analyses available to D&I researchers.80 For example, quantitative methods rely on existing survey instruments and the validity (or lack thereof) of those instruments.64,82 Therefore, quantitative methods are not useful for identifying new barriers/facilitators or exploring new constructs related to successful implementation if they are not included in existing measures or models. Similarly, quantitative methods are not well suited to provide explanations for novel or null findings—as they also are not represented in the hypotheses being tested, based on existing conceptual frameworks and models.

Quantitative methods emphasize breadth and generalizability of findings, and therefore they are also limited in their ability to provide depth of understanding for how and why implementation varies across different circumstances and settings.65,83 This is especially true when the subgroup of interest is small.64,66,84 Few quantitative methods are well suited to examine quantitative phenomenon with low variability or in small samples. For example, contrary to their theory-based hypotheses, Shapiro et al.64 found no significant associations between the perceived barriers/facilitators for program use and number of families served by the intervention. They hypothesized that this null finding may be in part due to lack of variability in the number of families served across providers (the majority served fewer than 10 families in the past year). However, they also pointed out that a small number of providers (N = 4) used the program with a very high number of families (more than 100 families in the past year). Because of the size of this subgroup, they were unable to statistically examine what facilitated this high level of program use. Fortunately, as discussed in the following section, many of the limitations associated with quantitative methods can be addressed with qualitative methods.

WHAT QUALITATIVE METHODS CAN AND CANNOT DO IN IMPLEMENTATION SCIENCE

“A salient strength of qualitative research is its focus on the contexts and meaning of human lives and experiences for the purpose of inductive or theory-development-driven research. It is a systematic and rigorous form of inquiry that uses methods of data collection such as in-depth interviews, ethnographic observation, and review of documents. Qualitative data help researchers understand processes, especially those that emerge over time, provide detailed information about setting or context, and emphasize the voices of participants through quotes. Qualitative methods facilitate the collection of data when measures do not exist and provide a depth of understanding of concepts.”20(p. 4)

As with quantitative methods, there are certain things that can and cannot be done with qualitative methods in D&I science. One of the most frequently cited reasons for using qualitative methods in implementation research is to understand barriers and facilitators to implementation29–31,42,46,85,86 and sustainment.45,87 While the findings of these studies point to several commonly occurring barriers and facilitators, thus increasing the generalizability of these qualitative findings, other studies have highlighted barriers and facilitators that are context or setting specific. For instance, in using data from focus groups and individual semistructured interviews, Fox and colleagues68 ascertained barriers and facilitators of implementing an evidence-based quality improvement intervention that have been identified in settings with other interventions (e.g., scheduling trainings for busy providers) and barriers and facilitators that were specific to both the setting and the intervention (e.g., audio-visual equipment that sometimes failed, space not being optimal at some sites, and complexities inherent in adapting a training initially developed for individual completion to a group setting).
Another reason for using qualitative methods is to document implementation processes. Patton argues for the use of qualitative methods for implementation evaluation because such evaluation "requires case data rich with the details of program content and context. Because it is impossible to anticipate in advance how programs will adapt to local conditions, needs, and interests, it is impossible to anticipate what standardized quantiles could be used to capture the essence of each program’s implementation. Under these evaluation conditions, a strategy of naturalistic inquiry is particularly appropriate.”

Aarons and Palinkas collected qualitative data through annual interviews and focus groups to assess the process of implementation of Safe Care, an intervention designed to reduce child neglect and out-of-home placements of neglected children into foster care. Lessard et al. used semistructured interviews with implementation facilitators to understand their role and the dynamics of facilitations between facilitators, family medicine groups, and other change actors. Hoagwood and colleagues used a case study of an individual child to describe the process of implementation of an evidence-based, trauma-focused, cognitive-behavioral therapy for treatment of symptoms of PTSD in children living in New York City in the aftermath of the World Trade Center attack on September 11, 2001. Related to an understanding of process is the use of qualitative methods to evaluate implementation success by identifying and explaining which aspects of the program are working or not working, for whom and in what circumstances and to understand how interventions are sustained.

As Albright et al. observe, “qualitative approaches are also useful when seeking to understand why evidence-based practices were successfully or unsuccessfully implemented, or when seeking strategies for facilitating implementation.”

Qualitative methods are used in implementation research to confirm or validate quantitative analyses through the technique of triangulation or convergence. In a pragmatic study comparing a nurse-administered tobacco tactics intervention to usual care, Duffy and colleagues collected and analyzed quantitative data from recruitment logs, chart audits, and surveys administered to study participants, as well as qualitative data collected from structured interviews. “The qualitative results were triangulated with quantitative results, which involved cross-verifying the same information from different sources, in this case comparing patient data, nurse survey and interview data, and EMR data.”

Use of qualitative methods to confirm or validate quantitative analyses is especially important in implementation research because the unit of analysis is the organization at which implementation occurs, resulting in sample sizes that are often too small to provide adequate statistical power. Saldana is using qualitative data to increase understanding of how behavior in agencies implementing evidence-based practices, as defined by the quantitative Stages of Implementation Completion (SIC) measure, relates to on-the-ground decision-making. In this instance, qualitative data is used to determine extent of convergence between reliable patterns of behaviors and particular patterns of SIC scores, “thereby helping to unpack the potential for SIC scores to serve as a proxy for less easily observed phenomenon.”

Qualitative methods are also well suited to understanding the context in which implementation occurs. “Understanding the context under which interventions work and how different ways of implementing an intervention lead to successful outcomes are required for “T3” (i.e., D&I of evidence-based interventions) and “T4” translations (i.e., policy development to encourage evidence-based intervention use among various stakeholders)

Qualitative methods can also be used to develop and establish construct validity of quantitative implementation outcome measures. According to Proctor and colleagues, “Qualitative data, reflecting language used by various stakeholders as they think and talk about implementation processes, is important for validating implementation outcome constructs.”

Martinez and colleagues recommend the use of mixed methods for addressing the instrumentation issue in implementation science of choosing the most appropriate evaluation method and approach.

In each of these instances, qualitative methods are used to elicit the perspectives of implementation stakeholders, including administrators, providers, and patients/clients.

“As Proctor and colleagues observe, “Any effort to implement change in care involves a range of stakeholders, including the treatment developers who design and test the effectiveness of ESTs, policy makers who design and pay for the service, administrators who shape program direction, providers and...”
supervisors, patients/clients/consumers and their family members, and interested community members and advocates. The success of efforts to implement evidence-based treatment may rest on their congruence with the preferences and priorities of those who shape, deliver, and participate in care.\(^\text{97}^\text{(p. 72)}\)

Elsey and colleagues\(^\text{95}\) used semistructured interviews with providers and patients to understand patient and health worker knowledge about tobacco and patients’ motivation to quit to inform the design of a cessation intervention. Kane and colleagues\(^5\) also conducted interviews with providers to assess perspectives on feasibility and acceptability of new guidelines.

While qualitative methods have important roles to play in the scientific endeavor that extends from identification of a phenomenon to active intervention, they are also limited in certain respects. For instance, while qualitative methods are often used to generate hypotheses related to D&I, they are rarely used for hypothesis testing. Some researchers have argued that qualitative data can be used to test hypotheses.\(^96\) Deductive approaches to analyzing the content of qualitative data where themes are identified a priori can be viewed as one form of hypothesis testing. Colon-Emeric et al.\(^\text{97}\) analyzed interview transcripts using framework analysis of a priori concepts, combined with inductive analyses. Sommerbakk and colleagues\(^\text{98}\) used a combination of thematic analysis using an inductive approach and theoretical thematic approach, applying codes to Grol & Wensing’s\(^\text{99}\) multilevel model of barriers and facilitators. Still other studies have relied on theoretical models to inform use of qualitative methods in mixed method implementation studies. Elsey and colleagues\(^\text{95}\) used Normalization Process Theory\(^\text{100}\) to understand barriers and facilitators to implementation of a tobacco cessation intervention in primary care settings in Nepal. However, other researchers have noted that qualitative methods are better suited to inductively generating hypotheses than deductively testing them.\(^\text{100}\) Still others take a “pragmatic approach,” arguing that inductive and deductive techniques can be employed in the same study in an iterative fashion. As Patton observes,

“The extent to which a qualitative study is inductive or deductive varies along a continuum. As evaluation fieldwork begins, an evaluator may be open to whatever emerges from the data, a discovery, or inductive approach. Then, as the inquiry reveals patterns and major dimensions of interest, the evaluator will begin to focus on verifying and elucidating what appears to be emerging — a more deductively oriented approach to data collection and analysis.”\(^\text{100}(p. 253)\)

Second, qualitative methods are generally not used to produce generalizable findings due to the lack of samples that are selected at random from a larger population and of insufficient size to provide adequate power for statistical analysis. Occasionally, efforts are made to insure that the information gained from one sample or study is “transferable” to another sample or setting, which is not the same as claiming that the findings obtained from a sample of a population is “generalizable” to other members of the population not sampled.\(^\text{100}\) However, although some forms of purposive sampling are designed to identify a range of variation in participant or organizational characteristics and behavior,\(^\text{101}\) such methods are not designed to eliminate known or unknown potential sources of bias that may limit generalizability. Nevertheless, as Padgett asserts, findings can have generalizability and resonance without being “generalizable” in a statistical sense bas on how the sample was selected.\(^\text{100}(p. 183)\)

Third, when quantifying and analyzing qualitative data using statistical methods, caution must be exercised in adhering to the assumption associated with the conduct of such analyses. It must be remembered that despite their numerical form, findings from the analysis quantitized qualitative data are primarily exploratory and descriptive.

**CASE STUDY**

**Background**

Several models of research translation have been proposed over the years (see Tabak et al.\(^\text{102}\) and Damschroeder et al.\(^\text{11}\) for a reviews). Many of these models consider sustainment to be the final stage of the process of implementation,\(^\text{103,104}\) but the factors that predict sustainment are not well understood\(^\text{10,105}\). In part, this may be attributed to a lack of consensus as to what constitutes sustainment and how to measure it. There are no uniform or agreed upon criteria for determining whether something has been sustained or not.\(^\text{106}\) This may be due to the fact that what is to be sustained differs from one program to the next. For instance, with respect to the community
coalitions supporting drug and suicide prevention activities, some definitions of sustainment focus on the coalition itself while others focus on the activities and impacts of the coalition. Furthermore, with few exceptions, most studies reporting positive results have focused on earlier stages of implementation progress (exploration, adoption, routine use) and not on sustainment. In addition to uncertainty as to how to define sustainment, there is a lack of consensus as to how to measure it.

Palinkas and colleagues describe a protocol for the development and evaluation of a system of measuring sustainment of US Substance Abuse and Mental Health Services Administration (SAMHSA)-supported prevention programs that target substance abuse prevention at the state or single community level, suicide prevention, and prevention of aggressive/disruptive behavior in elementary schools. An examination of all four grant programs simultaneously provides an opportunity to determine what is meant by the term “sustainment” and identify and support both the unique contextual requirements for improving sustainability for each program as well as for developing a generalizable framework comprised of core components of sustainment across diverse prevention approaches.

Context
The mixed method design is sequential, giving priority to the development of the quantitative measurement system (qual → QUAN) through convergence (comparing qualitative and quantitative data and quantizing qualitative data) and development (using qualitative data to develop a quantitative measure) by merging and connecting the data. In the first phase of the study, researchers interviewed 45 representatives of 10 grantees within four SAMHSA programs (Strategic Prevention Framework—State Initiative Grants, Sober Truth on Preventing Underage Drinking [STOP-Act], Garrett Lee Smith Suicide Prevention Program, and Prevention Practices in Schools). Data collection consisted of a semistructured interview to identify experiences with implementation and sustainment barriers and facilitators, a free listing exercise to elicit participant conceptions of the word “sustainment” and what it will take to sustain their programs; and a checklist of elements of the Consolidated Framework for Implementation Research to identify which are important for sustainment. Semistructured interview data were analyzed using a grounded theory approach of coding, consensus, co-occurrence and comparison. Coding was conducted by five members of the research team using the cloud-based Dedoose qualitative software. Comparisons of coding of the same transcripts by each team member revealed an average of 91.25% agreement across three levels of codes. Individual codes were then used to construct larger themes. Frequency counts and rank ordering of items obtained by the free list exercise were used to identify the salience of what is meant by sustainment/sustainability, program elements to be sustained, and requirements for sustainment. The procedure of constant comparison was used to identify meaningful clusters of items representing similar constructs. Implementation domains perceived to be relevant to sustainment were quantified by calculating the percent of informants citing the domain as high or very high in importance and a mean domain score (0 = not important, 1 = yes/no, 2 = important, 3 = very important). The lists of sustainment constructs obtained from the three sets of qualitative data were then compared side by side to identify those constructs appearing in more than one data set.

In the second phase of this project, these results along with an assessment on SAMHSA data collection systems will be used to develop a flexible measurement system, with both general and specific components, that can bring precision to monitoring sustainment of infrastructure, activities, and outcomes for each prevention approach. In the third phase, the sustainment measurement system will be administered to all current and recently grantees funded by the four SAMHSA programs to quantitatively evaluate its utility in monitoring progress and improving the likelihood of sustainment. This project is intended to improve sustainment of the supporting prevention infrastructure, activities, and outcomes that are funded by federal, state, community, and foundation sources.

Lessons Learned
Four elements for measurement were identified by all three data sets collected in Phase I of the project (ongoing coalitions, collaborations, and networks; infrastructure and capacity to support sustainment; ongoing evaluation of performance and outcomes; and availability of funding and resources), and five elements were identified by two of three data sets (community need, community buy-in and support, leadership, presence of a champion, and evidence of positive
outcomes). Some differences in construct priorities were also observed across programs, including norms; values and guiding principles of organizations; pressure from other states, tribes, or communities; perception of the current situation as needing change; knowledge and beliefs about the program; coalitions; collaboration and networking; and community specific activities that grantees wished to see sustained.¹¹²

**FUTURE DIRECTIONS IN MIXED METHODS**

As with any field of inquiry, scientific advancements often dictate or capitalize on advancements in methodology. This principle also applies to the use of mixed methods in D&I science. One potential focus of innovation in mixed methods lies in the development of a comprehensive understanding of implementation that links process to outcomes and both process and outcomes to context. Understanding the process of D&I is believed to be critical to understanding its outcome, but new causal models linking the two may require broader application of specific mixed methods such as qualitative comparative analysis.² While qualitative methods are appropriately used to gain an in-depth understanding of the context in which D&I occurs, some generalization of process and outcomes is necessary to achieve a level of understanding that extends beyond merely a collection of unique experiences and circumstances. Further efforts are required to identify the “sweet spot” that exists between the generalizable and the specific characteristics of D&I processes and outcomes.

A second focus of mixed methods innovation is the identification of new strategies to support implementation of innovative evidence-based programs and practices.²⁹ For instance, Palinkas and colleagues examined the role of social networks in scaling up evidence-based practices in county-level service systems in California and Ohio. Analysis of quantitative data of social networks found an association between the size and structure of social networks of child welfare, mental health, and juvenile justice systems leaders and stage of implementation of an evidence-based practice for treatment of behavioral health problems experienced by youth in foster care. Social networks were constructed using quantitative data collected from a web-based survey and qualitative data collected from semistructured interviews.¹¹⁵ A complementary analysis of semistructured interview data also revealed that collaborations between organizations were viewed by systems leaders as critical to implementing EBPs because they provided opportunities to exchange information and advice regarding specific EBPs and to pool resources, especially in small, rural counties where agencies have limited resources on their own.¹¹⁶ Three sets of collaboration characteristics were identified: (1) characteristics of the collaboration itself, including focus (within or across counties), formality and frequency; (2) characteristics of the external environment, including availability of funding, county size, existence of clients in need of the services provided by each agency, and government mandates; and (3) characteristics of participating organizations, including a common language, common recognition of the problem to be addressed, common goals and values, a buy-in and commitment to innovation and change, policies and procedures designed to ensure accountability, existence of interpersonal relationships and social networks, an individual who could serve as a broker or advocate for the EBP, leadership that was supportive of the implementation, and participants possessing qualities of honesty, credibility, trust, and respect for others.

Another implementation strategy identified by mixed methods is the use of research evidence. Analysis of qualitative data obtained from semistructured interviews and a focus group with child welfare, mental health, and juvenile justice systems leaders informed the development of the quantitative Structured Interview for Evidence Use (SIEU).¹¹⁷ The scale was then used to assess use of research evidence was collected from 151 directors and senior administrators of child welfare, mental health, and juvenile justice systems in 36 California and 9 Ohio counties participating in an RCT of the use of community development teams (CDTs) to scale up implementation of Treatment Foster Care Oregon (formerly known as Multidimensional Treatment Foster Care) over a 3-year period (2010–2012). Separate multivariate models were used to assess independent effects of evidence acquisition (input), evaluation (process), application (output), and total engagement (SIEU Total) on two measures of EBP implementation: highest stage reached and proportion of activities completed at preimplementation, implementation, and sustainment phases. Stage of implementation and proportion of activities completed in the implementation and sustainment phases were independently associated with acquisition of evidence and total engagement in
use of research evidence. Participation in CDTs was significantly associated with acquisition and total engagement in evidence use.118

SUMMARY

In reviewing the rapidly growing literature on the use of mixed methods to address important issues confronting the science of D&I, we conclude with two observations. The first observation is that mixed methods reflect an iterative process of data collection and analysis that involves both inductive and deductive approaches to understanding complex phenomena. As such, researchers may be forced to alter or abandon a priori strategies for data collection and analysis. How these activities actually occur may bear little resemblance to how they were imagined to occur when the grant application was prepared and submitted. This is especially true during the exploratory phases of a multiphase project, where determining the most appropriate means for collecting and analyzing data may lead to some form of methodological trial and error.

Second, this chapter began with the observation that using mixed methods is more than parallel play involving separate quantitative and qualitative studies. One of the implications of the defining characteristic that the methods must somehow be integrated is that the standards for insuring the rigor and appropriateness of each method when conducted as part of a mixed method strategy may not be the same as the standards when conducted independently. In a mixed method design, it is conceivable if not necessarily desirable that qualitative data may be analyzed quantitatively despite their failure to adhere to the assumptions of sufficient sample size, normality, and generalizability required for use of statistical tests. Such a practice may run counter to the disciplinary traditions of both quantitative and qualitative methodologists. The point here is that mixed methods represent both naturalistic inquiry and experimentation. The nature of D&I requires innovation in use of both quantitative and qualitative methods, and not every innovation will be successful.

Both the iterative nature of D&I science and the likely debate and compromises involved in selection and application of quantitative and qualitative methods in a mixed method design demand attention on the part of the investigators to document and detail the rationale for the selection of methods and the process and outcomes of their use. Without such documentation, the strengths and weaknesses of mixed method designs will be as context specific with limited generalizability and utility as the phenomena of D&I to which these methods are applied.

ACKNOWLEDGMENTS

Support for this chapter was provided by the following grant funded by the National Institute on Drug Abuse: R34 DA037516-01A1.

SUGGESTED READINGS AND WEBSITES

Readings


This article applies the typology for mixed method studies developed by Palinkas and colleagues to examining factors impacting staff retention during an evidence-based intervention implementation in a statewide child welfare system. The authors integrate qualitative data with previously published quantitative analyses of job autonomy and staff turnover in order to illustrate the utility of mixed method approaches in providing a more comprehensive understanding of opportunities and challenges in implementation research.


This article discusses a number of dimensions of mixed methods research, utilizing at least one qualitative method and at least one quantitative method, that may be helpful when designing projects or preparing grant proposals for conducting pragmatic trials and dissemination and implementation research.


This report was commissioned by the NIH Office of Behavioral and Social Science Research to serve as a resource that would provide guidance to NIH investigators on how to rigorously develop and evaluate mixed methods research applications. The report summarizes the nature and design of mixed methods research; teamwork, infrastructure, resources, and training for mixed methods research, incorporating mixed methods research in research, fellowship,
The Certificate in Mixed Methods Research is designed for researchers and practitioners in social work, nursing, psychology, public health, anthropology, political science, sociology, education, and other applied fields who are interested in: ways to integrate qualitative and quantitative research methods and data, commonly used qualitative and quantitative data collection methods and procedures, popular data analysis techniques used in the applied professions, and effective approaches to research conducted in practice settings.


The Mixed Methods Research Training Program for the Health Sciences is funded by the National Institutes of Health through the Office of Behavioral and Social Science Research (OBSSR), and is the only program of its kind in the United States. The program fulfills a national need for training in mixed methods and is a natural next step following the publication of the OBSSR “Best Practices for Mixed Methods Research in the Health Sciences”


The RAND Center for Qualitative and Mixed Methods (C-QAMM) develops and promotes tools for generating empirically based insights through iterative, exploratory data collection and analysis.

REFERENCES


Mixed Methods Evaluation in Dissemination and Implementation Science


60. Aarons GA, Green AE, Trott E, et al. The roles of system and organizational leadership in
system-wide evidence-based intervention sustain-
62. Green AE, Trott E, Willing CE, Finn NK, Ehrhart MG, Aarons GA. The role of collaborations in sustai-

63. Hunter SB, Han B, Slaughter ME, Godley SH, Garner BR. Associations between implementation charac-
65. Berkel C, Murray VM, Roulston KJ, Brody GH. Understanding the art and science of implementa-
67. Azocar F, Cuffel B, Goldman W, McCarter L. The impact of evidence-based guideline dissemination for the assessment and treatment of major depres-

sion in a managed behavioral health care organiza-
68. Beidas RS, Kendall PC. Training Therapists in Evidence-Based Practice: a critical review of stud-
69. Kilbourne AM, Abraham KM, Goodrich DE, et al. Cluster randomized adaptive implementation trial comparing a standard versus enhanced implementa-
71. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: concep-
72. Atkins MS, Frazier SL, Leathers SJ, et al. Teacher key opinion leaders and mental health consult-
76. Yost J, Ganann R, Thompson D, et al. The effective-
78. Davies P, Walker A, Grimshaw J. A systematic review of the use of theory in the design of guide-
line dissemination and implementation strategies and interpretation of the results of rigorous evalu-
79. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a sys-
tematic review and meta-analysis of the impact of theoretical basis, use of behavior change tech-
81. Chaudoir SR, Dugan AG, Barr CH. Measuring factors affecting implementation of health innov-
ations: a systematic review of structural, organi-
83. Gibbs DA, Krieger KE, Cutbush SL, Clinton-
Sherrod AM, Miller S. Implementer-initiated


