Children and Youth Services Review 105 (2019) 104425

Contents lists available at ScienceDirect

Children and Youth Services Review

journal homepage: www.elsevier.com/locate/childyouth

The group care quality standards assessment: A framework for assessment, quality improvement, and effectiveness

Shamra Boel-Studt[a,*], Jonathan C. Huefner[b], Hui Huang[c]

[a] College of Social Work, Florida State University, 296 Champions Way, Tallahassee, FL 32306, United States of America
[b] Boys Town National Research Institute, 14100 Crawford Street, Boys Town, Crawford, NE 68010, United States of America
[c] Robert Stempel College of Public Health & Social Work, Florida International University, 111200 SW 8th Street AHC5, Miami, FL 33199, United States of America

ARTICLE INFO

Keywords:
Residential group care
Therapeutic residential care
Child welfare
Quality improvement
Quality standards
Performance measures

ABSTRACT

Concerns about the quality of residential care for children and youth are longstanding. These concerns prompted a Florida-based initiative aimed at transforming residential care through the integration of research-informed practice standards, on-going assessment, and continuous quality improvement. The initiative resulted in the development of the Group Care Quality Standards and the Group Care Quality Standards Assessment (GCQSA) as mechanisms for guiding transformation efforts. In this article, we elaborate on the conceptual and implementation frameworks guiding the development and efforts to scale up the GCQSA throughout Florida. We begin by summarizing empirical sources that informed the guiding frameworks. Next, we describe the project phases highlighting the aims, methods and summarizing results where relevant. The aim of this article is to offer a working blue print to guide the adaptation of quality initiatives in other child welfare organizations or jurisdictions while taking into consideration the fit of such initiatives within the service environment and the complexities of system-wide change.

1. Introduction

Recently, there has been an increase in public concern about quality of care in residential programs for children and youth (Farmer, Murray, Ballentine, Rauktis, & Burns, 2017; James, Thompson, & Ringle, 2017; Pavkov, Negash, Lourie, & Hug, 2010). This is not a new concern, as the need for providing quality residential youth care has long been discussed (e.g., Allen, 1948; Healy & Bronner, 1937). Research findings support that quality services are related to better outcomes (Farmer, Seifert, Wagner, Burns, & Murray, 2017; Grietens & Hellinckx, 2004; Lee, Bright, Svoboda, Fakunmoju, & Barth, 2011; Whittaker et al., 2016). Quality social services have been defined as "the degree to which interventions influence client outcomes in desired ways in applicable domains while being delivered in a sensitive manner consistent with ethical standards of practice and the best available practice knowledge" (Megivern et al., 2007, p. 118).

Researchers have proposed establishing quality standards for residential care for children and adolescents (Boel-Studt & Tobia, 2016; Farmer, Murray, et al., 2017; Lee & McMillen, 2008). Federal guidelines, such as the Adoption and Safe Families Act of 1997 (ASFA; "Adoption Assistance and Child Welfare Act,", 1980) and the Family First Prevention Services Act (FFPSA; "Family First Prevention Services Act 2017,") have placed child well-being at the center of this debate (Wulczyn, Barth, Yuan, Harden, & Landsverk, 2017). For example, FFPSA requires that children are cared for in "a setting providing high-quality residential care" (section 472(k)(2)(D)). Clearly, no child should be referred to *any type* of service provision that struggles with maintaining safety and producing positive outcomes.

Two major efforts to articulate the elements of quality of residential services for children and adolescents occurred in the 1990s. The first was a commission in Scotland to establish a set of quality standards for use in residential child care. The task was to base this work on published knowledge and clinical expertise. The resulting quality standards were based on a literature review of factors related to positive outcomes for youth in residential care and extensive interviews with youth who had been in residential programs and their families. The report included 66 recommendations for improving the quality of care (Skinner, 1992). The second effort proposed a set of performance standards for residential care programs in the United States. The argument for these standards was that only programs offering high quality services could adequately meet the needs of children and adolescents with emotional and behavioral challenges. The proposed standards were measurable and directly related to child well-being. Ultimately, the developers argued that program performance based on quality standards would

* Corresponding author at: College of Social Work, Florida State University, 296 Champions Way, Tallahassee, FL 32306, United States of America.
E-mail addresses: sboelstudt@fsu.edu (S. Boel-Studt), Jonathan.Huefner@boystown.org (J.C. Huefner), huanhu@fiu.edu (H. Huang).

https://doi.org/10.1016/j.childyouth.2019.104425
Received 5 April 2019; Received in revised form 19 July 2019; Accepted 20 July 2019
Available online 23 July 2019
0190-7409/

facilitate continuous quality improvement within organizations and allow consumers to select the best programs for troubled children and adolescents. The published report proposed 20 standards, each with two to four related measures of quality (Daly & Peter, 1996). Neither of these early efforts, however, have been widely adopted or implemented. Consequently, there continued to be a felt need to establish quality standards for residential care for children and adolescents (James, 2017; McMillen et al., 2005; Zayas, McMillen, Lee, & Books, 2013)

In 2015, an opportunity to promote quality in residential care programs arose in Florida. Florida was experiencing pressure to reduce or eliminate its highest levels of care. In response, a provider network in the state advocating on behalf of residential group care for children and adolescents (RGC) effectively shifted the discourse from reduction-elimination to a need for quality improvement, laying the groundwork for a statewide initiative aimed at establishing and implementing quality standards. Florida's Group Care Quality Standards Initiative is a collaboration between the Florida Department of Children and Families (DCF), the Florida Institute for Child Welfare (FICW), the Florida Coalition for Children (FCC), academic researchers, child advocates, and service providers and consumers aimed at improving the quality and effectiveness of RGC.[1] The initiative draws upon research and empirically-driven frameworks to transform residential services for children and adolescents through the integration of research-informed practice standards, on-going assessment, and continuous quality improvement.

In a previous article Boel-Studt, Huefner, Bender, Huang, & Abell, 2018, we described the background of the initiative and development of the *Group Care Quality Standards* (Group Care Quality Standards Workgroup, 2015) and *Group Care Quality Standards Assessment* (GCQSA). In this article, we elaborate on the conceptual and implementation frameworks guiding development and efforts to scale up the GCQSA statewide. We begin by summarizing empirical sources that informed the guiding frameworks. Next, we describe the project phases highlighting the aims, methods and summarizing results where relevant. The aim of this article is to offer a working blue print to guide the adaptation of quality initiatives in other child welfare organizations or jurisdictions while taking into consideration the fit of such initiatives within the service environment and the complexities of system-wide change.

2. Overview of empirical sources

The conceptual framework of GCQSA draws upon a vast literature with primary sources including Megivern et al.' (2007) model of influences on quality social services, implementation science (Bertram, Blase, & Fixsen, 2015; Fixsen, Blase, Naoom, & Wallace, 2009; Ghate, 2016) and principals of developmental evaluation (Patton, 2010).

2.1. Model of influences on quality social services

Drawing upon Megivern et al. (2007) definition of quality social services, we further define quality standards as specified, operational conditions guiding the provision of quality care. Importantly, this encompasses technically proficient practices based on knowledge and proper use of effective methods and an approach to service delivery that is ethical and sensitive to clients' individuality and culture. Building upon quality assurance literature, Megivern et al. (2007) present an integrated model of quality social services outlining both the multidimensional components of quality care and factors that influence service quality. The model depicts structural (organizational receptivity and capacity), processes (proficient technical care, provider receptivity, interpersonal and cultural sensitivity), and system and consumer outcomes (sustained functioning, reduction in problems, subjective well-being) as quality elements and influences. It further accounts for the role of macrosystem engagement, advocacy, consumer engagement, and the overarching influence of stakeholder feedback as conditions that influence quality care. These elements represent the real-world complexities that determine quality care with implications for informing quality improvement. Many of these components informed the conceptual and implementation frameworks underlying the GCQSA highlighting key considerations for effective and sustainable implementation within programs and system wide.

2.2. Implementation science and practice

Multiple promising frameworks have evolved from the field of implementation science and practice to guide implementation of evidence-based or innovative practices within agencies or whole systems (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005; Ghate, 2016). A number of these frameworks outline stages of implementation moving from conceptualization and planning to full implementation and sustainability (Hanson, Self-Brown, Rostad, & Jackson, 2016). Although the elements of the frameworks vary, common stages generally include exploration/planning, installation, initial implementation, full implementation, and sustainability (Albers, Mildon, Lyon, & Sholonsky, 2017; Hanson et al., 2016; Nilsen, 2015). During the exploration phase, key stakeholders are engaged along with the needs, fit, and capacity of the system to accommodate a new practice. Implementation drivers, (i.e., the capacity and infrastructural supports necessary to successfully implement new practices) are considered and key partners and implementation teams established. Instillation involves bringing together resources needed to support implementation. Moving to initial implementation, the new practice is introduced into the practice setting. Evaluative processes support further development and refinement of implementation strategies. Once the new practice is embedded into the system, full implementation is considered to have been achieved when 50% of providers are routinely using the new practice with fidelity (National Implementation Research Network, 2017). However, on-going evaluation and efforts to support fidelity and sustainability remain in place. The implementation stages evolve in a non-linear process of on-going planning, learning, evaluation, and development.

2.3. Developmental evaluation

Recognizing the dynamic and, oftentimes, unpredictable nature of the child welfare system and service environment, the developmental evaluation approach (Patton, 2010), that embraces flexibility in implementation as a means of achieving ecologically informed and valid programs, builds in on-going evaluation, creating stakeholder feedback loops, to inform an iterative development process. This approach is fitting for the development of innovative practices recognizing that formulating an effective practice is a continual process essentially informed by application in the actual service environment. In distinguishing developmental evaluation from traditional formative and summative evaluation approaches that embody a process of "ready, aim, fire" to inform program processes linked to outcomes, developmental evaluation takes an orientation of "ready, fire, aim" and repeat (as cited in Patton, 2010, p. 30). This highlights a bottom up, continuous approach to program development; recognizing the new practice as a dynamic evaluative target. Moving further along in development, more traditional, formative and summative evaluation can be built-in at appropriate phases.

[1] The Florida Department of Children and Families is the state child welfare agency in Florida. The Florida Institute for Child Welfare is a state funded research institution housed at the Florida State University College of Social Work. The Florida Coalition for Children is a non-profit organization comprised of a network of child service providers in Florida who advocate on behalf of maltreated children and youth.

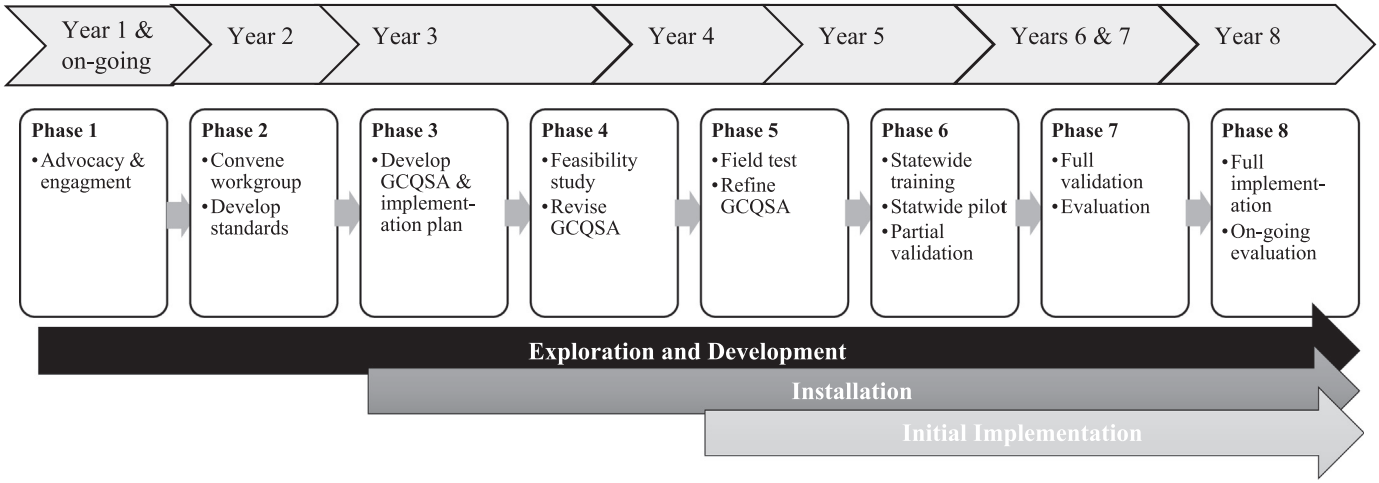

Fig. 1. Florida quality standards initiative implementation framework.

3. GCQSA conceptual and implementation frameworks

Considering the literature and guiding frameworks, the approach of the project team is one of on-going learning and development, informed by data-driven processes and the expertise of a diverse stakeholder group. Fig. 1 summarizes the eight-phased implementation plan guiding development and the process of scaling up the GCQSA and the proximal alignment with the frameworks of implementation science and practice. The project timeline begins with the initial advocacy efforts through full system-wide integration. Sustainability efforts are weaved in throughout the phases. The figure depicts phases that have already occurred (phases 1–5), are in-progress (phase 6), and planned (phases 7–8). In the following sections, we discuss the purpose and primary activities of each phase.

3.1. Advocacy and engagement (Phase 1)

In 2014, the viability of RGC in Florida was at the forefront of discussions among policy makers, fueled by negative media and advocacy initiatives aimed at "right-sizing" (i.e., greatly reducing) or "eliminating" RGC. Legislative proposals along these lines were presented despite a lack of empirical support and without sufficient evidence examining the potential impacts of reducing or eliminating group care, including stressing an already overburdened system by eliminating placement options and exacerbating existing challenges with finding stable, effective placements for youth with higher level needs. Leaders of the Florida Coalition for Children (FCC), a statewide provider network, and the Florida DCF came together to discuss the role of RGC on the child welfare service continuum. Three RGC provider members of the FCC advocated to balance the information on group care being disseminated throughout the state. They engaged legislators, lead agency administrators,[2] residential group care providers statewide, and the media. In their messaging, they acknowledged quality issues while presenting data and research demonstrating quality RGC programs can produce positive youth outcomes. These efforts were successful in shifting the discourse from reduction/elimination and instrumental in laying the groundwork for Florida's Quality Standards Initiative. Crucial to the effectiveness of these initial efforts was the active engagement of stakeholders from multiple levels to promote a common message, creating a platform for change. Two key results of these efforts were the forming of the Group Care Quality Standards Workgroup and the eventual passing of state legislation aimed at quality accountability for children's residential group homes. Nearly two years into the project, continued advocacy efforts by the DCF and workgroup leaders lead to the passing of H.B. 1121 during Florida's 2017 legislative session. The work of the project team along with a successful feasibility study and on-going implementation pilot provided strong impetus for passing the bill that includes the following language:

> The Department shall develop, in collaboration with the Florida Institute for Child Welfare, lead agencies, service providers, current and former foster children placed in residential group care and other community stakeholders, a statewide accountability system based on measurable quality standards (FL HB 1121, "Child Welfare Act,", 2017, p. 82).

The quality standards developed by the workgroup and the GCQSA serve as the foundation of the mandated statewide accountability system for RGC.

3.2. Group Care Quality Standards (Phase 2)

The DCF in partnership with the FCC convened the Group Care Quality Standards Workgroup, comprised of 26 members including group care providers and child advocates throughout Florida with research support provided by the FICW and Boys Town National Research Institute. The workgroup included providers from a broad range of residential care facilities throughout Florida. The workgroup was tasked with developing a set of research-informed quality standards that were considered applicable to all licensed residential group homes. Huefner's (2018) consensus of proposed practice standards for RGC provided the workgroup with a working list of standards grounded in empirical research and best practice guidelines. In his review, Huenfer synthesized seven multinational sources to identify sixty-four quality standards, organized into eight practice domains. Lead by FCC Residential Committee leadership, members of the workgroup divided into task groups assigned to discuss the proposed standards within one of the eight practice domains to select and adapt standards for Florida's group homes. The standards identified by the task groups were reviewed and compiled into one document, resulting in the published guide, *Quality Standards for Group Care* (Group Care Quality Standards Workgroup, 2015). The guide outlines a set of 59 quality practice standards in the following domains: 1) Assessment, Admission, and Service Planning, 2) Positive, Safe Living Environment, 3) Monitor and Report Problems, 4) Family, Culture, and Spirituality, 5) Professional and Competent Staff, 6) Program Elements, 7) Education, Skills, and

[2] Florida's child welfare system began implementing a privatized system in 1999. The state uses a Community-Based Care model, in which services are contracted with local agencies who are responsible for coordinating services and providing case management.

Positive Outcomes, and 8) Pre-Discharge/Post Discharge Processes.

3.3. Group Care Quality Standards Assessment (Phase 3)

In phase three, the project team began work on the GCQSA designed to operationalize and measure group homes' implementation of the quality standards. Our aims were to design a mechanism for measuring quality to inform and facilitate a process of continuous quality improvement. Adding to the project team, we engaged former foster youth previously in group care, child advocates, and child welfare scholars to serve as reviewers of the various iterations of the GCQSA. Our objectives were to design and validate an assessment that captured the core, measurable elements of the quality standards defined by the workgroup and a process for implementing the assessment within RGC programs statewide. Key activities related to exploration and installation were completed during this phase including considering the fit of the assessment within the existing system, identifying available and needed resources and implementation drivers necessary to inform, activate, and oversee implementation. During this phase, we delineated the conceptual and implementation frameworks underlying the GCQSA.

3.3.1. GCQSA conceptual framework

The six elements of the conceptual framework include: 1) research-informed practice, 2) quality enhancement, 3) multi-informant 4) multidimensional 5) equitable, and 6) reliable and valid. Beginning with *research-informed practice,* a specific aim of the initiative is to establish and measure standards grounded in research. To achieve this, the team drew upon an evidence-based practice process, combining available research synthesized in Huefner's (2018) publication and the expertise of multiple stakeholders. Infused in the quality standards are empirically supported elements of effective RGC and best practices informed by a consensus process among field experts and consumers. The *quality enhancement* orientation, is reflected in the decision to align the standards with state licensing standards and to implement the assessment as part of the annual re-licensure process. The assessed quality standards expand upon the minimum state licensing standards, which primarily address structural elements of quality; assessing whether programs have the capacity to provide an appropriate care setting. The quality standards expand upon and add to these criteria while shifting the assessment approach to examine the extent to which quality care is provided from a multidimensional perspective. That is, the state licensing standards serve as minimum quality requirements while the standards assessed in the GCQSA represent quality enhancements grounded in research and best practices. For example, the state licensing code requires programs to complete an assessment for each child to determine admission and service planning. The quality standards expand on this by requiring the use of a validated, comprehensive assessment that specifically assesses safety risks, child strengths, needs, and prior traumas. For a detailed description of the licensing review that was completed as part of the development of the GCQSA readers are referred to Boel-Studt et al. (2018).

The GCQSA is a *multi-informant* assessment comprised of four forms that measure the standards from the perspectives of different stakeholders. This includes three separate self-report forms completed by youth, group care providers (i.e., directors and director care workers), and lead contract agency staff (e.g., case managers, contract managers, placement coordinators). The fourth form is completed by DCF licensing specialists[3] through a combination of document reviews, site observations, and youth and staff interviews. For all forms, respondents are instructed to provide a rating of how consistent practices or conditions in the group home are with a given statement using a 5-point scale ranging from “Not at all” to “Completely”. Mean scores are based on the combined ratings across stakeholders to provide a global assessment for each domain. The *multidimensional* definition of quality is captured within the eight practice domains identified in the Group Care Quality Standards (2015) and the eight corresponding scales of the GCQSA. The items measuring the specific standards are partitioned into structural, process (Donabedian, 1988, 2002) and experiential measures (Morris & Bailey, 2014) representing different attributes of quality. Structural items measure the infrastructure of the care setting such as staffing, policies, facilities, and resources (e.g., GCQSA Licensing Form: There are a minimum of two staff on duty at all times, except during sleeping hours.). Process items measure the extent to which providers consistently follow recommended service guidelines (e.g., GCQSA Provider Form: Youth are involved in creating their service plans.) while experiential items assess consumers' experiences within the care environment (e.g., GCQSA Youth Form: I am allowed to do regular things that most kids do, like play sports, spend time with friends, or go to school events.). The standards encompass the technically proficient aspects of care through assessing congruence with best practice, research and policy-based guidelines. Other items are designed to tap the “interpersonal and culturally sensitive aspects of care” (Megivern et al., 2007, p. 117); assessing consumer experiences, use of approaches that maximize cultural awareness and consumer inclusiveness, voice, and empowerment wherever possible (e.g., GCQSA Provider Form: Youths' spiritual beliefs and values are supported through providing access to places where they can practice their beliefs (e.g., churches, temples, mosques)).

Another key consideration is ensuring the assessment is *equitable*; targeted to the aspects of the care environment that are within the providers' capacity to control or influence. For instance, involving youths' families in care must be carefully assessed recognizing that the courts, case managers, and families all influence the extent to which family involvement is achieved. Ensuring equitability is adequately captured required a sufficient amount of piloting and on-going evaluation of the fit and relevance of the standards within the real-world practice context that encompasses a heterogeneous population of group homes operating within complex internal and external service environments. Understanding these aspects will yield a more accurate measure and inform how the results should be interpreted and utilized. Importantly, it also serves as an opportunity to contextualize how policies and practices that are part of the larger system influence quality of care and, given these realities, what can be considered realistic expectations of RGC providers. Relatedly, this iterative process of fitting the standards to the context is instrumental in establishing and implementing ecologically relevant standards and *reliable and valid* measures. We believe this is the crux of successful quality initiatives that hold true potential to impact practice. The assessment was designed to allow for rigorous analyses of its psychometric properties (i.e., reliability and validity) which are being tested through a series of pilots and a validation study. Ecological validity, a form of external validity, is described as the ‘real-life representativeness’ of findings generated from research (Magill, 2011). In the present initiative, we view ecological validity as reflective of the applicability and practicability of the standards in practice from the standpoint of those living and working in the practice setting. Evaluative components designed to assess these aspects were integrated into the project phases that began with engaging stakeholders in the process of defining the standards and developing the GCQSA. As part of the pilot testing of the GCQSA, evaluative components included focus-group style discussion sessions with stakeholders and including text response items on the GCQSA forms to elicit stakeholder feedback. These data helped refine the how the standards are assessed.

3.3.2. GCQSA implementation framework

The GCQSA is not only a measure of RGC quality. It is a system-wide

[3] Licensing specialists are employees of the Department of Children and Families who are responsible for conducting annual site inspections of all residential homes to ensure conditions are in compliance with the state licensing code.

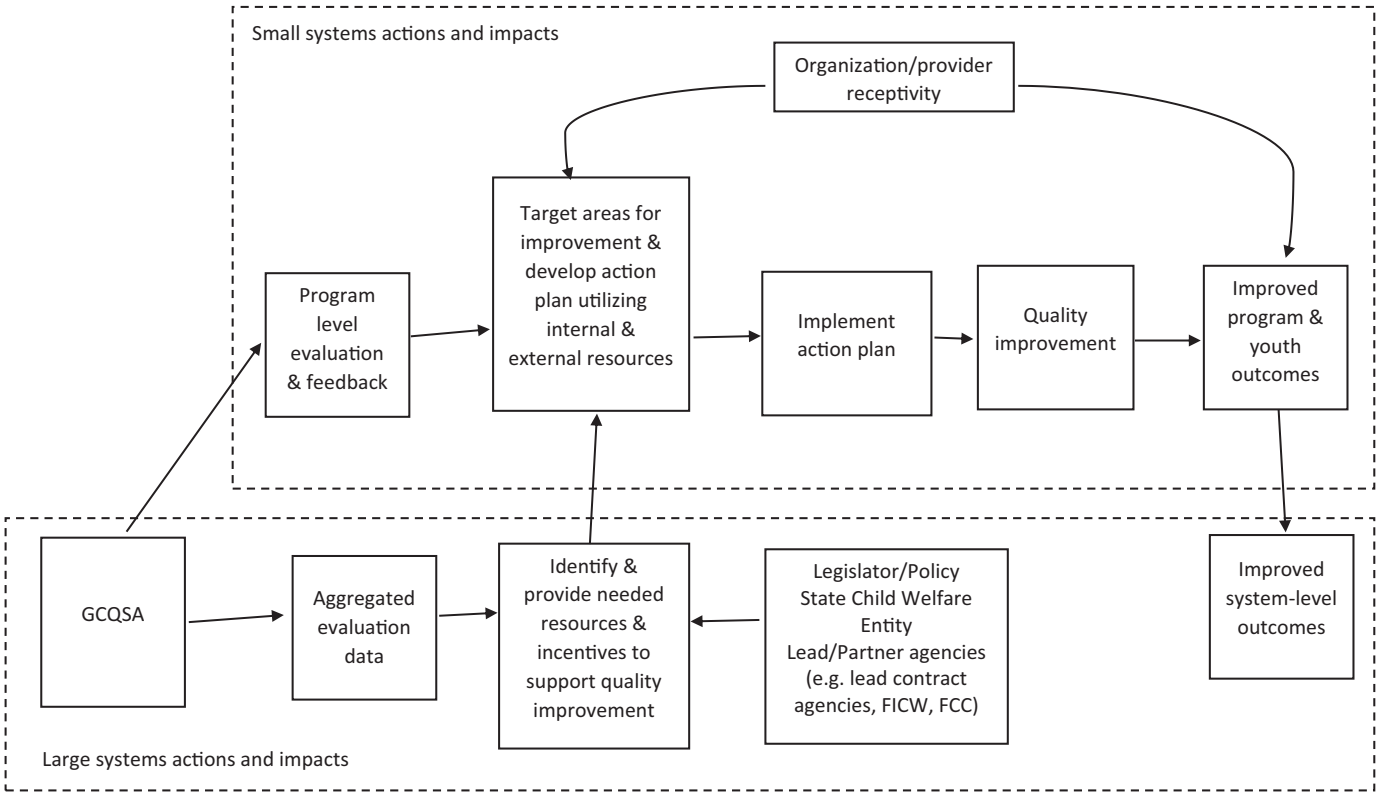

Fig. 2. Group care quality standards theory of change.

intervention. Fig. 2 depicts the processes moving from assessment to utilizing the results to inform quality improvement planning and implementation, ultimately, contributing to program and system-wide improvements. Notably, this requires identifying a set of program and system-level outcomes. Beginning with the licensing teams, the GCQSA is completed for each RGC program as an added-on component of the annual re-licensure inspection. This creates a process of on-going evaluation and feedback loops. Positioning the assessment within state licensing promotes the uptake of a universal set of standards while making quality assessment accessible to all RGC providers by eliminating costs. At the program level, providers can use assessment results to identify areas to target for improvement that informs actions plans and benchmarking to track progress. At the larger-system level, aggregated evaluation data can be utilized by lead agencies in the counties, regions, or state to target resources to facilitate quality improvements. These collective efforts have potential to yield program improvements and, subsequently, quality improvement and positive outcomes system-wide. The effectiveness of these efforts are maximized with strong organizational and provider receptivity and external supports (e.g., training/technical support, financial incentives, political environment).

3.4. Feasibility study (Phase 4)

In phase four, we completed a feasibility study to evaluate the implementation plan (for details see Boel-Studt et al., 2018). Questions guiding the feasibility study focused on how well the implementation plan fit with the existing re-licensure process and assessing stakeholder participation and views on the process and assessment items. The study sample included 11 group homes located in central Florida that were due for their annual re-licensure inspections during the two-month study period. Data were collected using a pilot version of the GCQSA. Implementation process data were collected from a combination of text response items on the GCQSA requesting participant feedback, technical support calls with the regional licensing team, and an in-person debriefing session with providers at the completion of the study.

A total of 94 assessment forms were completed (56 service provider forms, 27 youth forms, 11 licensing specialist forms). Response rates ranged from 100% among group home directors, direct care workers, youth, licensing specialists, and case managers to 70% among placement coordinators. Overall, participants expressed support for the assessment and felt it was “lengthy” but “manageable”. We found the licensing team was able to accommodate the added tasks of providing oversight of the assessment and completing their own assessment forms. The biggest issue to emerge was related to the sampling procedures and how to best sample homes with different operational structures. For the feasibility study, one assessment was completed for every licensed residential facility. The population of group homes in the region and state varies widely ranging from small, private single-family homes to large agency-run institutions with multiple locations. We found variations in licensing where for some group homes with multiple sites each individual home was separately licensed whereas for others a license many encompass two or more homes which created challenges with the sampling protocol. The sampling issue was threefold. On the one hand, it was important that expectations and the target of the assessment were consistent as possible across group homes. Relatedly, questions centered on ensuring the sampling method resulted in assessment data that was representative of quality within the group homes and the extent to which ratings from one facility, even those operated by the same agency, could be meaningfully generalized to another. These concerns extended to determining which stakeholders should participate in the assessment and how to approach that for the varying operational structures. In short, the sampling method needed to yield an assessment that was reasonably representative of quality in a group home, of the participants who are affected by and/or affect the quality of care in group homes, and that was feasible to implement to ensure reliable data collection.

The majority of feedback reflected that participants viewed the

standards as "reasonable" and "relevant". Results of a preliminary reliability analysis of the youth and provider forms were promising for the overall scales (Youth $\alpha = 0.90$; Provider $\alpha = 0.73$). We found youth and external providers (e.g., case managers, placement coordinators) were unable to provide informed ratings of items from the *Professional and Competent Staff* and *Monitor and Report Problems* scales. This was reflected in a high proportion of "Not Applicable' ratings ($\geq$ 90%) and confirmed in follow-up discussions with licensing specialists and providers during the debriefing session. The items focused on staff training, supervision, and internal policies and procedures of the group home. Moving forward these scales were removed from the youth form and revised for the external provider form. The overall results demonstrated the feasibility of implementing the GCQSA within the state licensing system while providing data needed to adapt procedures to improve the fit within the service setting and to further adapt items and scales for different respondent types.

3.5. Implementation pilot (Phase 5)

The purpose of the implementation pilot was to evaluate the assessment and implementation protocol using a larger sample of group homes located in Florida's Central and Northeast regions, which have different service environments and licensing teams. The GCQSA was completed for 37 group homes (Central = 55.9%; Northeast 44.1%) by licensing specialists (37 forms), youth (72 forms), case managers and placement coordinators (78 forms) and group home directors and direct care workers (116 forms). Consistent with the feasibility study, completion rates remained high across respondents ranging from 79 to 100%. Cronbach's alpha was calculated for all eight scales for all four forms (Youth $\alpha = 0.60$–0.90, Licensing $\alpha = 0.32$–0.86, Group Care Provider $\alpha = 0.46$–0.81, Contract Agency $\alpha = 0.73$–0.94). Cronbach's alphas for two scales were below acceptable for the *Monitor & Report Problems* scale ($\alpha = 0.46$) of the group home directors/direct care worker form and *Professional and Competent Staff* scale ($\alpha = 0.32$) of licensing specialist form.

Focus-group style debriefing sessions with participants ($n = 24$) focused on assessing the implementation process and participants' perspectives on the applicability and practicability of the standards. Thematic analyses (Braun & Clarke, 2006) revealed themes related to implementation indicating few issues with participation across respondents and a continued need to further adapt the sampling methods to increase fit within the practice context. Participants shared mixed views on the applicability and practicability of some of the standards. Their responses suggested that some standards reflected practices that were currently not the normative practice across group homes and would likely require a substantial shift in perspectives and practices (e.g., family involvement in treatment). Some participants expressed mixed views related to practicability where some believed certain standards reflected areas that were the responsibility of the case manager (e.g., working with psychiatrists to manage medications) while others reported engaging in these practices. This divergence in views and practices underscores the importance of taking into account providers' interpretations and views related to the standards as this likely translates into why standards may or may not be followed and have implications for interpreting quality ratings. These discussions also pointed to areas where the standards may have potential to positively affect practice through facilitating incremental changes and provided guidance for revising items to more precisely capture aspects of care that are within the group home provider's range of influence (equity) and understanding.

The combined results from the feasibility study and implementation pilot demonstrated that a promising framework for the assessment had been established and provided critical insights to guide measurement revisions and successful integration across the state. Challenges with the sampling methods encountered in both studies resulted in further adaptations so that one assessment could be completed for each group care program. That is, each group care program, whether the program consists of a single group home or multiple homes located on one campus or with different physical addresses within a region, could be represented by one assessment as long as the homes followed the same practice model. This approach represented a compromise between ideal conditions of external validity and what we found was feasible/practical in the service setting.

3.6. Statewide Pilot Study (Phase 6)

The statewide pilot study represents a major step toward full implementation. The purpose of the statewide pilot is to begin implementing the GCQSA in all six regions, giving participants in each region an opportunity to become familiar with the assessment while providing careful monitoring and on-going technical assistance. The statewide pilot includes all DCF licensed group homes and shelters throughout the state representing > 150 RGC programs. The pilot began with a day-long orientation and training held in each of Florida's six regions. Training sessions included group care providers, lead agency personnel, and the regional licensing teams. Applying similar methods as in the two previous pilots, we are collecting GCQSA data that will be used to finalize the assessment in preparation for validation. We are also collecting process data via technical support calls and survey data to evaluate implementation that will inform final adjustments to facilitate accuracy, efficiency, and fidelity to the process.

A major role of the project lead team (FICW and DCF) throughout the statewide pilot has been providing technical support, especially for the licensing teams to facilitate form completion and troubleshoot issues that arise. Another objective has been to collect data on the population of group homes throughout the state that will be used to create a comprehensive index of RGC programs outlining program structures, models, services, and service populations. Building in the statewide pilot prior to formal validation allows the team to identify and address implementation challenges that may otherwise interfere with data collection and data quality. However, some steps toward validation described in the next section may be carried out during this phase provided these challenges are manageable and will not impact data quality.

3.7. Validation Study (Phase 7)

The purpose of the validation study is to examine the psychometric characteristics of the GCQSA. Using research methods established in the previous pilots, the team will collect assessment data for the full population of group homes over the course of one-year to coordinate with the existing re-licensure timeline. Analyses focus on scale reliability and validity proceeding in stages that are generally sequential but not neccessarily linear. That is, results revealed in one anlaysis may prompt a need to return to a prior analysis for further consideration or modifications. Reliability or the extent to which a measure performs consistently across repeated uses as evidenced by inter-item correlations can be evaluated using Cronbach's alpha. Confirmatory factor analysis (CFA) can be used to examine scale dimensionality, specifically, in the Florida version of quality assessment, to test the hypothesized eight factor structure of the GCQSA. Item response theory (IRT; see also latent trait theory) is an increasingly preferred method of measurement validation. IRT analyses provide information to assess item performance (e.g. difficulty, discrimination, and information). Additionally, multidimensional item response theory (MIRT) models can be used to examine dimensionality and model fit (Cai, 2010; Muthén & Asparouhov, 2013). Although preferred, due to a larger number of parameters being estimated, IRT models require larger samples which are not always feasible, especially in applied research. Both traditional (e.g., CFA) and IRT methods can be applied to guide item selection and assess scale properties. Osteen (2010) recommended integrating both methods when feasible to maximize unique information obtained from

each.

During this phase, we will also examine measurement validity or the extent to which the scale (GCQSA) measures the construct (quality RGC) it was designed to measure. Different types of validity exist, however, in the preliminary stages of instrument validation, establishing construct validity is prioritized (Abell, Springer, & Kamata, 2009). Construct validity is partitioned into convergent validity and discriminant validity. Convergent validity is established by examining correlations between the new scale being validated and another scale or variable thought to measure a similar construct (evidence of convergent relationships). If the scores are found to be correlated, evidence is provided that the new scale measures the intended construct. Conversely, discriminant validity is established by comparing scores from the new scale with scores of a disimilar measure wherein a lack of correlation supports that the new scale is measuring a distinct construct from that of the disilimiar scale (evidence of discriminant relationships). Ideally, standardized scales can be used to test convergent validity. However, single-item indictors can also provide evividence of convergent validity when standardized instruments do not exist as in the case of the GCQSA. To our knowledge, this is the first measure of its kind to be validated. Single-item indictors represent “direct, straightforward definitions of core constructs being validated” (Abell et al., 2009, p. 67). Ensuring that these measurement characteristics have been achieved to an acceptable degree is necessary prior to applying the results of the GCQSA to assess quality of care.

Once these intial validation steps are complete, interest in testing other forms of reliability and validity may be worth pursuing. In the present study, for instance, criterion validity focusing on whether quality ratings predict a select set of program or youth outcomes (i.e., predictive ‘criterion’ validity) is of interest. Efforts to select outcomes and, particularly, to pose hypotheses regarding predictive validity of scale scores should be carefully wieghed in full consideration of the pros and cons of varying results. Additionally, inter-rater reliability, which reflects consistency of ratings bewteen two more raters assessing the same unit of anlaysis (e.g., group home) using the same assessment (e.g., GCQSA) is another possibility. Results of inter-rater reliability analyses can add support for the reliability of an assessment but also yeild useful information when ratings are found to differ between raters that can inform training those completing the assessment to increase reliability. That is, to ensure the ratings are a reflection of the construct being assessed and not of the rater.

3.8. Full Implementation (Phase 8)

The final phase marks full implementation. Decisions about scoring and interpretation of the GCQSA will need to be finalized along with reporting procedures and mechanisms for monitoring continued implementation. Baseline data on group homes' performance on the standards can be collected following validation with continued efforts to monitor trends over time reflecting the extent to which the standards are intregrated into practice. The selection of performance outcomes at the program, region, and/or state level can provide further evidence of the impact of the quality standards. Finally, because the evidence-base for RGC continues to evolve and grow, the quality standards should be periodically reviewed and updated to reflect the current state of evidence-informed/best practice knowledge.

4. Discussion

The Florida Group Care Quality Standards initiaive aims to address gaps in the evidence-base and practice of RGC through establishing and measuring a validated set of quality performance standards that can be used to guide quality improvements. Experience has shown that licensure and accreditation, while important, and have not been sufficient to ensure quality. Additionally, accreditation is costly and typically limited to larger programs. The GCQSA brings in-depth quality focus to all residential programs. In this article, we describe the guiding frameworks underlying the design and efforts to scale up the GCQSA statewide. Beginning with advocacy efforts, we decribe key phases and activies within each phase facilitating a process that is consistently grounded in an evidence-based practice process, data driven, and inclusive of key stakeholders at multiple levels ranging from providers and youth to child advocates and state legislators.

Our goal is to provide a general framework that can be useful to facilitate quality standards and improvement efforts in RGC programs beyond Florida. The Florida Group Care Quality Standards initiaive can serve as a case study offering a blue print to leaders in other jurisdictions who are looking for quality measures and an example of how such measures can be implemented. Although a substantial amount of literature has shown that quality of group care is associated with youth outcome (Farmer, Murray, et al., 2017; Grietens & Hellinckx, 2004; Lee et al., 2011; Whittaker et al., 2016), limited literature documents how research findings are translated into practice. This article contributes to literature by sharing our lessons learned about how to translate research findings into an assessment tool for quality assurance practice.

Designed to be transportable, the standards and guiding frameworks can be adapted to fit multiple service contexts. We believe certain elements such as engaging key stakeholders, taking an incremental approach to scale development, validation, and implementation including sufficient piloting, attention to ecological validity along with rigorous validation and clear goals and objectives are key to effective quality initiatives. To every extent possible, this work was grounded in existing research and empirical frameworks along with innovative thinking that comes with implementing a new process and a shared vision and commitment among the project team. Ultimately, the project team envisions a child welfare service continuum that can offer high quality and effective services for children and youth with high level care needs for whom lower levels of care have not proven to be sufficient.

Now in its fifth year, this initative continues to make steady progress; meeting the objectives outlined in each of the phases. Efforts to measure and explore the impacts of the quality standards initiative on program and youth outcomes are underway. However, preliminary evidence of early impacts on practice are emerging. The licensing specialists report that providers are using the GCQSA to evaluate their program models and are moving toward greater implementation of assessments to guide service planning. The continuity and success of this initiative thus far are largely due to a committed network of stakeholders and successful advocacy. All to often innovative ideas and projects lose momentum part-way through due to external forces or other influences resulting in wasted resources and missed opportunities. It's important that external forces not cause those invested in such initiatives to lose site of the key issues that, regardless of idealogical or policy shifts, continue to exist. Our hope is that the initative will receive continued support and that others will join us by engaging in similar efforts to transform RGC through promoting service quality and effectiveness. These efforts are paramount to detemining RGC's appropriate place on the child welfare service contiuum.

Acknowledgements

Funding

This study is funded by the Florida Institute for Child Welfare (Grant No. 038750) and the Florida Department of Children and Families (Grant No. 036331).

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Declaration of Competing Interest

None of the authors of this paper has competing interests, financial or otherwise, in any of the findings of this study.

References

Abell, N., Springer, D. W., & Kamata, A. (2009). *Developing and validating rapid assessment instruments.* New York, NY: Oxford University Press.

Adoption Assistance and Child Welfare Act (1980). *42 U.S.C. 1305 et seq*.

Albers, B., Mildon, R., Lyon, A. R., & Sholonsky, A. (2017). Implementation frameworks in child, youth and family services–results from a scoping review. *Children and Youth Services Review, 81*, 101–116. https://doi.org/10.1016/j.childyouth.2017.07.003.

Allen, F. H. (1948). Developments in child psychiatry in the United States. *American Journal of Public Health, 38*(9), 1201–1209.

Bertram, R., Blase, K. A., & Fixsen, D. L. (2015). Improving programs and outcomes: Implementation frameworks and organization change. *Research on Social Work Practice, 25*(4), 477–487. https://doi.org/10.1177/1049731514537687.

Boel-Studt, S. M., & Tobia, L. (2016). A review of trends, research, and recommendations for strengthening the evidence-base and quality of residential group care. *Residential Treatment for Children & Youth, 33*(1), 13–55. https://doi.org/10.1080/0886571X.2016.1175995.

Boel-Studt, S., Huefner, J. C., Bender, K., Huang, H., & Abell, N. (2018). Developing quality standards and performance measures for residential group care: Translating theory to practice. *Residential Treatment for Children & Youth,* 1–22. https://doi.org/10.1080/0886571X.2018.1536494.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa.

Cai, L. (2010). A two-tier full-information item factor analysis model with applications. *Psychometrika, 75*(4), 581–612. https://doi.org/10.1007/s11336-010-9178-0.

Child welfare Act of 2017 (2017). FL HB 1121 s.409.996(22). Retrieved from https://www.flsenate.gov/Session/Bill/2017/1121/BillText/er/PDF.

Daly, D. L., & Peter, V. J. (1996). *National performance standards for residential care: A policy initiative from father Flanagan's' Boy's home.* Boys Town, NE: Father Flanagan's Boy's Home.

Donabedian, A. (1988). Quality assessment and assurance: Unity of purpose, diversity of means. *Inquiry, 25*(1), 173–192.

Donabedian, A. (2002). *An introduction to quality assurance in health care.* New York, NY: Oxford University Press.

Family First Prevention Services Act of 2017, H.R. 253 2017.

Farmer, E. M. Z., Murray, M. L., Ballentine, K., Rauktis, M. E., & Burns, B. J. (2017). Would we know it if we saw it? Assessing quality of care in group homes for youth. *Journal of Emotional and Behavioral Disorders, 25*(1), 28–36. https://doi.org/10.1177/1063426616687363.

Farmer, E. M. Z., Seifert, H., Wagner, H. R., Burns, B. J., & Murray, M. (2017). Does model matter? Examining change across time for youth in group homes. *Journal of Emotional and Behavioral Disorders, 25*(2), 119–128. https://doi.org/10.1177/1063426616630520.

Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice, 19*(5), 531–540. https://doi.org/10.1177/1049731509335549.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). Core implementation components. *Implementation research: A synthesis of the literature* (pp. 24–34). Tampa, FL: University of South Florida.

Ghate, D. (2016). From programs to systems: Deploying implementation science and practice for sustained real world effectiveness in services for children and families. *Journal of Clinical Child & Adolescent Psychology, 45*(6), 812–826. https://doi.org/10.1080/15374416.2015.1077449.

Grietens, H., & Hellinckx, W. (2004). Evaluating effects of residential treatment for juvenile offenders by statistical metaanalysis: A review. *Aggression and Violent Behavior, 9*(4), 401–415.

Group Care Quality Standards Workgroup (2015). *Quality standards for group care.* Boys Town, NE: Boys Town Press. http://www.dcf.state.fl.us/programs/foster-care/group-care/docs/1507-255%20Quality%20Standards%20for%20Residential%20Group%20Care%20Booklet.pdf.

Hanson, R. F., Self-Brown, S., Rostad, W. L., & Jackson, M. C. (2016). The what, when, and why of implementation frameworks for evidence-based practices in child welfare and child mental health service systems. *Child Abuse & Neglect, 53*, 51–63. https://doi.org/10.1016/j.chiabu.2015.09.014.

Healy, W., & Bronner, A. F. (1937). New light on delinquency and its treatment. *The Howard Journal of Crime and Justice, 4*(4), 391–397.

Huefner, J. C. (2018). Crosswalk of published quality standards for residential care for children and adolescents. *Children and Youth Services Review.*. https://doi.org/10.1016/j.childyouth.2018.03.022.

James, S. (2017). Implementing evidence-based practice in residential care: How far have we come? *Residential Treatment for Children & Youth, 34*(2), 155–175. https://doi.org/10.1080/0886571X.2017.1332330.

James, S., Thompson, R. W., & Ringle, J. L. (2017). The implementation of evidence-based practices in residential care: Outcomes, processes, and barriers. *Journal of Emotional and Behavioral Disorders, 25*(1), 4–18. https://doi.org/10.1177/1063426616687.

Lee, B. R., Bright, C. L., Svoboda, D. V., Fakunmoju, S., & Barth, R. P. (2011). Outcomes of group care for youth: A review of comparative studies. *Research on Social Work Practice, 21*(2), 177–189.

Lee, B. R., & McMillen, J. C. (2008). Measuring quality in residential treatment for children and youth. *Residential Treatment for Children & Youth, 24*(1–2), 1–17. https://doi.org/10.1080/08865710802146622.

Magill, M. (2011). Moderators and mediators in social work research: Toward a more ecologically valid evidence base for practice. *Journal of Social Work, 11*(4), 387–401. https://doi.org/10.1177/1468017310379930.

McMillen, J. C., Proctor, E., Megivern, D., Woodstock Striley, C., Cabassa, L., Munson, M., & Dickey, B. (2005). Quality of care in the social services: Research agenda and methods. *Social Work Research, 29*(3), 181–191.

Megivern, D. M., McMillen, J. C., Proctor, E. K., Striley, C. L., Cabassa, L. J., & Munson, M. R. (2007). Quality of care: Expanding the social work dialogue. *Social Work, 52*(2), 115–124.

Morris, C., & Bailey, K. (2014). *Measuring health care quality: An overview of quality measures.* (Retrieved from Washington, DC: I:\national research institute\Electronic Articles\Morris&Bailey14.pdf).

Muthén, B. O., & Asparouhov, T. (2013). Multidimensional, multilevel, and multi-time-point item response. In W. J. van der linden (Ed.). *Handbook of item response theory: Models, statistical tools, and applications* (pp. 527–540). Boca Raton, FL: CRC Press.

National Implementation Research Network (2017). *Implementation stages.* Chapel Hill, NC: University of North Carolina. Retrieved from https://nirn.fpg.unc.edu/learn-implementation/implementation-stages.

Nilsen, P. (2015). Making sense of implementation theories, models and frameworks. *Implementation Science, 10*(1), 53. https://doi.org/10.1186/s13012-015-0242-0.

Osteen, P. (2010). An introduction to using multidimensional item response theory to assess latent factor structures. *Journal of the Society for Social Work and Research, 1*(2), 66–82. https://doi.org/10.5243/jsswr.2010.6.

Patton, M. Q. (2010). *Developmental evaluation: Applying complexity concepts to enhance innovation and use.* New York, NY: Guilford Press.

Pavkov, T. W., Negash, S., Lourie, I. S., & Hug, R. W. (2010). Critical failures in a regional network of residential treatment facilities. *American Journal of Orthopsychiatry, 80*(2), 151–159.

Skinner, A. (1992). *Another kind of home: A review of residential child care.* Glasgow, SCT: The Scottish Office.

Whittaker, J. K., Holmes, L., del Valle, J. F., Ainsworth, F., Andreassen, T. H., Anglin, J. P., ... Zeira, A. (2016). Therapeutic residential care for children and youth: A consensus statement of the international work group on therapeutic residential care. *Residential Treatment for Children & Youth, 33*(2), 89–106. https://doi.org/10.1080/0886571X.2016.1215755.

Wulczyn, F., Barth, R. P., Yuan, Y.-Y. T., Harden, B. J., & Landsverk, J. (2017). *Beyond common sense: Child welfare, child well-being, and the evidence for policy reform.* New York, NY: Routledge.

Zayas, L. E., McMillen, J. C., Lee, M. Y., & Books, S. J. (2013). Challenges to quality assurance and improvement efforts in behavioral health organizations: A qualitative assessment. *Administration and Policy in Mental Health and Mental Health Services Research, 40*(3), 190–198.