Residential Group Care Accountability System
ANNUAL REPORT

Department of Children and Families
Office of Quality and Innovation
October 1, 2022

Shevaun L. Harris
Secretary

Ron DeSantis
Governor

Table of Contents

Purpose 3

Background 3

Quality Practice in Group Care

Oversight Activities 4

Quality Standards Assessment Tool

Accountability System

Conclusion 5

Appendix 6

PURPOSE

The Florida Department of Children and Families (Department) engaged the Florida Institute for Child Welfare (Institute) to develop and validate an assessment tool to measure, document, and facilitate quality services in Department licensed child-caring agencies, also referred to as group care. The Quality Standards for group care was established to set core quality standards to ensure that each program is managed equally to provide high-quality services to children in care.

Florida Statutes, section 409.996(25), requires the Department, in collaboration with the Institute, to develop a statewide accountability system for group care providers based on measurable quality standards. The accountability system is mandated to include the following:

1. Promote high quality in services and accommodations, differentiating between shift and family-style models, and programs and services for children with specialized or extraordinary needs, such as pregnant teens and children with the Department of Juvenile Justice involvement.

2. Include a quality measurement system with domains and clearly defined levels of quality. The system must measure the level of quality for each domain, using criteria that group care providers must meet to achieve each level of quality. Domains may include, but are not limited to, admissions, service planning, treatment planning, living environment, and program and service requirements. The system may also consider outcomes six months and 12 months after a child leaves the provider's care. However, the system may not assign a single summary rating to group care providers.

3. Consider the level of availability of trauma-informed care and mental health and physical health services, providers' engagement with the schools that children in their care attend, and opportunities for children's involvement in extracurricular activities.

BACKGROUND

The Group Care Quality Standards Workgroup was established in 2015 by the Department and the Florida Coalition for Children (FCC) to develop core quality standards for residential child-caring agencies (group homes) licensed by the Department. In addition, the Group Care Quality Standards Workgroup created the Quality Standards for Group Care to aid children in receiving high-quality services that surpass the minimum thresholds currently assessed through licensing. The workgroup was comprised of 26 stakeholders, including the Florida Institute for Child Welfare, group care providers, Community-Based Care Lead Agency staff, and other stakeholders. From the workgroup, a draft set of standards was developed and approved by the Department.

The approved quality standards are broken into the following eight domains:

Quality Practice in Group Care – Eight Domains

1. Assessment, Admission, and Service/Treatment Planning
2. Positive, Safe Living Environment
3. Monitor & Report Problems
4. Family, Culture, & Spirituality
5. Professional & Competent Staff
6. Program Elements
7. Education, Skills, & Positive Outcomes
8. Pre-Discharge/Post-Discharge Processes

The Department asked the Institute to take the lead on the development of a project plan that consisted of eight phases, including the following:

1. Advocacy and engagement
2. Development of core quality performance standards
3. Development of a quality assessment tool
4. Feasibility pilot
5. Implementation pilot
6. Statewide implementation
7. Full validation study and evaluation
8. Full implementation and ongoing evaluation

OVERSIGHT ACTIVITIES

Accountability System

During the 2021-2022 report year, the Department and the Institute completed the statewide validation study and the inter-rater reliability and agreement (IRRA) study. Data collection for the IRRA was completed in January 2021, and data collection for the statewide validation study was completed in March 2021. These components represent major steps toward fully validating the Group Care Quality Standard Assessment (GCQSA). A full description of both studies is provided in the subsequent report, along with detailed findings on the status of each and interim findings.

The Department will continue working with the Institute for the next two-year transition period of the project. A live webinar training was conducted in July 2022 on generating and interpreting QSA provider reports will be created to assist the licensing specialist. Additionally, a prerecorded training will be created during this year for new hires and boosters and ongoing technical assistance will be provided to support the licensing teams across the state.

QUALITY STANDARDS ASSESSMENT TOOL

With an approved set of quality standards and project plan, the Institute took the lead on the development and validation of an assessment tool designed to measure group providers within the eight domains. The GCQSA is comprised of four separate forms, which include: 1) Service Provider Form A, 2) Service Provider Form B, 3) Youth Form, and 4) Licensing Specialist Form. The assessment tool consists of three types of questions: structural, process, and experiential. Structural items measure the infrastructure of the group care setting (e.g., staffing, policies, resources), process items measure the extent to which providers consistently provide services that follow recommended guidelines, and experiential items measure experiences of consumers and providers within the group care setting. The Institute utilized an investigative approach to develop fully informed ratings for providers. These ratings were gathered through multiple sources to include document reviews, observations, interviews with program directors, staff and youth, experience, and judgment.

As a part of this effort, the Institute completed an extensive report titled, *An Assessment of Quality Standards for Florida's Department of Children and Families Licensed Residential Group Homes: Fiscal Year 2021-2022 Final Report*. This report provides a detailed description as to the following:

- Summary of Key Findings from the Statewide Pilot, Validation Study and Inter-rater Agreement Study.
- Outcomes Development Pilot.

- Quality Standards Assessment and Statewide Accountability System.
- Recommendations for sustainability.

See Appendix A for the full report, titled: *An Assessment of Quality Standards for Florida's Department of Children and Families Licensed Residential Group Homes: Fiscal Year 2021-2022 Final Report*

CONCLUSION

The Department continues to advance toward completion of the statutory requirements and goals associated with the Quality Standards for Group Homes contained in section 409.996, Florida Statutes.

The outcomes development pilot was an effort to measure change over time in youth placed in a child-caring agency. Despite challenges with a small sample, the results suggest higher quality ratings on select standards spanning most domains and were associated with improvement in youth behavioral and emotional difficulties.

The next action items for the Department and Institute include data collection to identify trends pertaining to the performance on the quality standards beginning July 1, 2022. In addition, the Institute will build provider reports and export reports to a project landing page where they can be accessed by Department staff and licensing specialists. The Institute will work with the Department data quality team on further report building (e.g., Dashboard).

Appendix A

An Assessment of Quality Standards for Florida's Department of Children and Families Licensed Residential Group Homes: Fiscal Year 2021-2022 Final Report

June 30, 2022

Research Report

June 30, 2022

Project Title: An Assessment of Quality Standards for Florida's Department of Children and Families Licensed Residential Group Homes FY21-22

Principal Investigator:

Shamra Boel-Studt, PhD, MSW
Florida State University College of Social Work

Project Team:

Taylor Dowdy-Hazlett, MSW
Florida State University College of Social Work

Hui Huang, PhD, MSSW
Florida International University, Robert Stempel College of Public Health and Social Work

Jonathan Huefner, PhD
Child and Family Translational Research Center, Boys Town

Contents

Executive Summary 2
Background 3
Summary of Key Findings from the Statewide Pilot, Validation Study and Inter-rater Agreement Study 3
Statewide pilot study 3
Statewide validation 5
Interrater agreement 6
Outcomes Development Pilot 7
Method 7
Measures 8
Description of data collected 8
Results 8
Description of Quality Standards Assessment and Statewide Accountability System 14
Conclusions 14
Recommendations for Sustainability 15

Executive Summary

Effective July 1, 2017, Section 409.996(25), Florida Statutes, required the Department of Children and Families (Department) to develop a statewide accountability system (SAS) for residential group care providers based on measurable quality standards. The Statute required the SAS to be fully implemented by July 1, 2022. Efforts leading up to the Statute included convening a statewide workgroup (Group Care Quality Standards Workgroup) tasked with establishing a set of research-based quality performance standards (Group Care Quality Standards Workgroup, 2015). The Department engaged the Florida Institute for Child Welfare (FICW) to lead the development and validation of the Quality Standards Assessment (QSA) to serve as the core measure of the SAS. The QSA was designed to measure residential program's performance on 59 quality standards across eight domains.

Development began in late 2015 and through a series of implementation pilots and studies, the QSA is fully developed, validated, and implemented statewide. In 2021, a pilot study was conducted to explore associations between the quality standards and youth's behavioral and emotional outcomes. The results comparing changes in youth behavioral and emotional symptoms identified marked improvements in youth and provider reported internalizing problems and total difficulties that were both clinically significant and statistically significant. The results of these exploratory analyses found 17 standards were associated with lower internalizing, conduct problems, and total difficulties at follow-up. That is, youth placed in programs with higher performance ratings on these standards experienced significantly greater improvements. Standards from seven out of eight domains were identified as significant. Most of the significant standards were from Doman 6. *Program Elements* (six standards) followed by Domain 1. *Assessment, Admission, & Service Planning* (four standards), Domain 7. *Education, Skills, & Positive Outcomes* (two standards) and Domain 2. *Safe, Positive Living Environment* (two standards). Three standards were associated with multiple outcomes – D6. *Quality Improvement Approach*, D6. *Actively Monitor Youth and Milieu,* and D6. *Psychotropic Medications Appropriately Monitored*.

The Department will continue to work with Dr. Boel-Studt and the FICW during a two-year transition period. During this time, Dr. Boel-Studt will provide an updated live webinar training with a component focusing on generating and interpreting QSA provider reports. A prerecorded training will be available for new hires and booster training purposes. Additionally, Dr. Boel-Studt will provide ongoing technical assistance to support the licensing teams in completing the QSA, build provider reports and export reports to a project landing page where they can be accessed by Department staff and licensing specialists. Dr. Boel-Studt will work with the Department data quality team on further report building (e.g., Dashboard) and any transition plans. Finally, trend data on programs' performance on the quality standards will be collected with baseline data collection beginning July 1, 2022. Recommendations for promoting maximum benefits of the QSA as part of the SAS, include focusing on defining and developing processes to systematically measure youth outcomes and to utilize QSA results to offer resources and training to support quality improvement efforts in Florida's residential group homes.

Background

Effective July 1, 2017, Section 409.996 (25), Florida Statutes, requires the Department of Children and Families (Department) to develop a statewide accountability system (SAS) for residential group care providers based on measurable quality standards. The Statute requires the SAS to be fully implemented by July 1, 2022. Efforts leading up to the Statute included convening a statewide workgroup (Group Care Quality Standards Workgroup) tasked with establishing a set of research-based quality performance standards (Group Care Quality Standards Workgroup, 2015). The Department engaged the Florida Institute for Child Welfare (FICW) to lead the development and validation of the Quality Standards Assessment (QSA),[1] which serves as the core measure of the SAS. The QSA was designed to measure residential program's performance on 59 quality standards across the following eight 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 Positive Outcomes
8. Pre-Discharge/Post Discharge Processes

Following initial regional pilots, a statewide pilot roll-out of the QSA and a validation study, the QSA has demonstrated evidence of reliability and validity. To date, the QSA represents the most rigorously developed and tested assessment of quality residential care. This report summarizes research aimed at validating the QSA, findings from pilot studies examining associations between the Quality Standards and program performance and outcomes, and efforts to finalize the QSA, SAS, and sustainability plans.

Summary of Key Findings from the Statewide Pilot, Validation Study, and Inter-rater Agreement Study

Efforts to validate the QSA began during development by establishing content validity (i.e., Do the items reflect the constructs they were designed to measure?) assessed by a panel of 16 subject matter experts (Boel-Studt et al., 2018). Elements of ecological validity (i.e., Do the concepts measured have real world applicability and practicability?) were evaluated during the feasibility study and implementation pilots (Boel-Studt & Huang, 2018). Initial estimates of internal consistency (i.e., Are scale items designed to measure the same constructs correlated across repeated uses?) were examined during these early phases to provide preliminary evidence of scale score reliability (Boel-Studt et al., 2018). Taken together, the findings from these earlier phases informed revisions to the QSA and implementation procedures, leading to the statewide pilot study.

STATEWIDE PILOT STUDY

[1] The previous title, *Group Care Quality Standards Assessment* (GCQSA), was shortened to the *Quality Standards Assessment* (QSA) for brevity and to include potential for broader applications in other service settings (e.g., youth shelters, residential treatment centers, juvenile justice centers, statewide in-patient psychiatric programs).

The purpose of the statewide pilot was to begin implementing the QSA across all six regions of the state (Boel-Studt, 2019). This began with a series of trainings with the licensing teams, residential care providers, and Lead Agencies in each region. After the training, QSA data were collected for one-year. The QSA was completed for each residential program as part of annual relicensure. Throughout the pilot, the project team provided technical assistance to the licensing teams and completed an evaluation of the implementation and performance of the QSA with a larger, statewide sample.

A total of 1,516 assessment forms were completed by youth (29.7%), direct care staff (29.7%), residential program directors (17.9%), Lead Agency staff (12.1%) and licensing specialists (10.6%), representing 160 residential programs. Regular documented technical assistance (TA) calls (N = 61) with the regional licensing teams were used to evaluate the implementation process, with few noted challenges. Exploratory and confirmatory factor analyses were used to examine the QSA factor structure for all five forms (youth, licensing, Lead Agency, director, direct care worker). After model re-specifications, all five forms demonstrated adequate fit (i.e., supported eight-factor scale structure for all four childcare professional forms and seven-factor scale structure for youth form). Reliability coefficients for domain scores across all forms were good to excellent (Youth α = .83-.92, Licensing Specialist α = .79-.94, Lead Agency Staff α = .88-.94; Residential Program Director α = .88-.96, Direct Care Staff α = .80-.91). Preliminary analyses using youth and licensing specialist QSA data were conducted to identify correlations between higher quality ratings and the occurrence of critical incidents (i.e., youth hospitalizations, staff injury, youth injury, law enforcement calls, runaway episodes, total incidents) reported by programs during a six-month period. Table 1 displays correlations between QSA Domain scores from the youth ratings and the number of critical incidents reported over a six-month period. Statistically significant, negative correlations were found for four out of five incident types. Further, all significant correlations (see bolded values notated with asterisks) were in the moderate range indicating higher quality ratings were associated with moderate decreases in the number of incidents reported by programs. Specifically, programs with higher quality performance ratings by youth in domains 1-4 and 6-7 reported significantly fewer youth injuries. Higher quality ratings across standards in all domains of the youth QSA were associated with significantly fewer police calls to campus and higher quality ratings in the Positive, Safe Living Environment and Program Elements domains were associated with moderate decreases in runaway episodes. The number of total incidents reported decreased as ratings on domains 1-2, 4 and 6 increased.

Table 1. Youth QSA Domain Score Associations with Six-Month Critical Incidents

QSA Domain	Critical Incidents					
	Hospitalized	Staff Injury	Youth Injury	Police calls	Runaway	Total Incidents
D1. Assessment, Admission, & Service Planning	-.14	-.05	**-.25****	**-.29****	-.14	**-.21***
D2. Positive, Safe Living Environment	-.11	.05	**-.31***	**-.22***	**-.23***	**-.21***
D3. Monitor & Report Problems	-.12	.03	**-.29***	**-.20***	-.11	-.16

D4. Family, Culture, & Spirituality	-.07	-.04	**-.23***	**-.28****	-.11	**-.22***
D6. Program Elements	-.15	.04	**-.22***	**-.40****	**-.23***	**-.31****
D7. Education, Skills, & Positive Outcomes	-.16	.01	**-.27****	**-.29****	-.12	**-.22***
D8. Pre & Post Discharge Planning	-.10	.10	-.17	**-.27***	-.19	**-.21***

**p<.05; **p<.01; Correlation coefficients between .1-2. are considered small, .2-.5 are considered medium, and >.5 are considered large. Note, Domain 5 Professional and Competent Staff is excluded from the youth QSA.*

Results of correlational analyses of the licensing form also showed mean scores on Education, Skills, and Positive Outcomes (r = -.23, p <.05) were associated with fewer law enforcement calls. No other domain scores on the licensing form were statistically associated with critical incidents. However, several item-level correlations were observed across most domains, demonstrating promising preliminary findings supporting associations between higher performance on the quality standards and fewer program incidents (for detailed results see Boel-Studt, 2019, Appendix B).

Themes from open-ended responses from licensing specialists provided insights into the ratings and current practice in residential programs related to family involvement, documentation, program models, and behavioral management. Emerging evidence suggested that participating in the QSA was prompting positive changes that may promote quality improvement. For example, licensing specialists reported that some residential providers were evaluating their models, developing plans for becoming trauma-informed, and improving documentation of training, supervision, and service plans. A content analysis of youths' (n = 119) open-ended responses indicated that overall, youth viewed their placements positively and felt connected with program staff. Youth also expressed concerns about environmental restrictions and having their needs adequately met by staff (for detailed results see Boel-Studt, 2019, pp. 18-19). Following additional modifications to the QSA, implementation plan and training, a statewide booster training was held via live webinar. With the additional refinements and training, data collection for the full validation study began in January 2020.

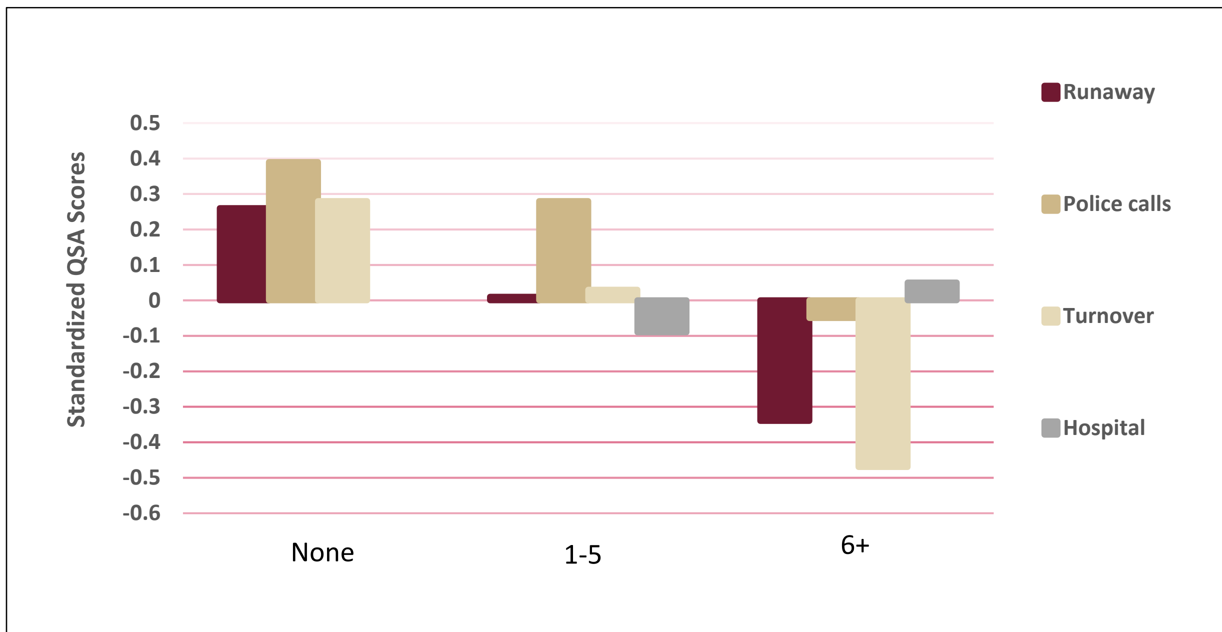
STATEWIDE VALIDATION

The purpose of the validation study was to examine construct validity of the QSA (i.e., does the QSA measure what it was designed to measure?). Convergent validity (a form of construct validity) is determined by comparing scores from the measure being tested against other measures or indicators designed to measure similar constructs. First, because no other validated measures of quality residential care exists, eight single item indicators (SII) designed to measure each of the quality standards domains were developed in consultation with measurement experts and practice experts. SSIs provide a concise, overall definition of the construct at the domain level. For example, Domain 1 "Assessment, Admission, and Service Planning" scores are measured based on mean ratings of the standards that comprise the domain. Drawing upon the standards that make up the domain, the corresponding Domain 1 single item indicator is "*Overall assessment, admissions, and service plans are completed with youth, parents/legal guardians, and other professionals, and consider strengths, safety, appropriate level of care and dignity of the youth and famil*y." The SII and QSA are both rated using the same 5-point scale (1 = not at all consistent with how things are done in this program, 5 =

completely consistent with how things are done in this program). Pearson's correlation coefficient was used to examine evidence of convergence between mean domain scores and domain SII ratings. For all five QSA forms, all correlations with the corresponding SII were positive, moderate to large (range .48-.94), and statistically significant, lending support for construct validity (for details see Boel-Studt et al., 2021, pp. 3-4, Table 2).

Next, QSA domains scores and reported program incidents, initially examined during the statewide pilot, were re-examined, this time using 12-month incident reports for each program and drawing upon global scores generated from standardized ratings from all five QSA forms (i.e., youth, licensing specialist, directors, direct care staff, Lead Agency staff; see Boel-Studt et al., 2021). Chart 1 displays a key finding showing a general trend where programs reporting a lower number of incidents (i.e., runaway, police calls, staff turnover) had higher QSA scores compared to programs with a higher number of reports. Taken together the results from the statewide validation study lend support for the construct validity of the QSA.

Chart 1. QSA Total Scores and 12-Month Incident Reports



Note. Critical incidents were separated into categories resulting in approximately equal groups and extreme outliers were dropped.

INTERRATER AGREEMENT

Interrater agreement (IRA) of the youth, Lead Agency, direct care workers, and residential director forms of the QSA were evaluated using data from the 2018-2019 statewide pilot study. To examine the IRA of the QSA licensing form, QSA data was collected from pairs of licensing specialists from five regions (Central, Northwest, Southeast, Southern, and Suncoast) who agreed to participate in the supplemental study. In each region, at least six group homes were assessed by two raters using the QSA. IRA data were available for 142 residential programs that were rated by two or more youth, 25 residential programs rated by two Lead Agency personnel, 131 programs rated by two or more direct care staff, 71 rated by two or more directors or supervisors, and 31 rated by two or more licensing specialists. The results (see Boel-Studt et al. 2021, Tables 5-6, p. 6) show that for forms completed by

direct care workers, directors, and Lead Agencies, Domains 2-7 showed acceptable IRAs, indicated by at least 60 percent of residential programs rated by pairs of respondents had moderate-high IRA with the values of rwg(j)[2] ranging from 0.50 and 1. For the youth form, Domains 2 and 6 showed satisfactory IRAs, indicated by at least 60 percent of residential programs demonstrating moderate-high IRA. For the other domains on the youth form, at least 50 percent of residential programs had moderate-high IRA. For the licensing form, all the domains showed acceptable IRAs. These results provide preliminary evidence of inter-rater agreement, an important form of reliability for assessments designed to be completed by multiple raters.

Outcomes Development Pilot

The purpose of the outcomes pilot was to explore associations between the quality standards and youth outcomes. Specifically, this study sought to explore 1) which standards are most strongly associated with youth outcomes, and 2) whether QSA scores distinguish between high and low performing programs, based on change in youth behavioral and emotional symptoms. Item-response theory (IRT) analyses are appropriate for responding to research question two and require a large sample size (e.g., 300-500). Due to sample size limitations, analyses for the second research question could not be performed at this time. The results presented focus on research question one. All study procedures were reviewed and approved by the Florida State University Institutional Review Board. Participation in this study was completely voluntary and no incentives were provided to participants.

METHOD

The study sample was identified using data from Florida Safe Families Network (FSFN), Florida's state child welfare database. FSFN data exports were provided to the Principal Investigator (PI) by the Department at three points over a six-month period. The first export included 1,606 youth placed in a residential program or shelter between February 25, 2020 and June 17, 2021. The second export included 430 youth placed in a residential program or shelter between June 21, 2021 and August 11, 2021, and the third export included 328 youth placed in a residential program or shelter between November 1, 2021 and December 10, 2021. The initial sampling frame, combining data from all three FSFN exports, was 2,466. Youth placed before June 1, 2021 (*N* = 1,395), under age 10 (*N* = 73) or who were placed in a shelter or Non-DCF licensed facility (*N* = 270), were excluded from the sample. The sample was further adjusted for duplicate entries for youth who changed residential placements during the six-month period (*N* =114). The adjusted sample was separated by the region of the youths' current residential placement (i.e., Central, Northeast, Northwest, Southern, Southeast, Suncoast). The lists of residential programs were sent to the regional licensing teams for further vetting (e.g., checking for licensing status, closures, non-DCF facilities), and to provide contact information for the programs where the sample of youth were currently placed. One program was eliminated due to closure and an additional 58 were dropped due to having no contact information provided. The final eligible sample included 555 youth from 103 residential programs.

Data collection included three phases: an initial assessment and two-follow ups. At Time 1, emails were sent to residential programs with study information and links to the *Strengths and Difficulties Questionnaire* (SDQ). For each youth, the youth and a program staff were asked to complete an SDQ. Following the initial email, 82 residential providers indicated the youth was no longer in their program.

---

[2] rwg(j) is an index of agreement for multi-item scales. For reference see James, L.R., Demaree, R.G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69*, 85-98.

These youth were dropped from the sample with no further follow up attempts made. The adjusted sample included 473 youth. Following the initial requests, three follow-up emails were sent to prompt completion of missing SDQs. Approximately 60 days (Time 2) following initial data collection, follow up emails were sent to providers of youth with an SDQ completed at Time 1 using the same procedures. A final follow-up (Time 3) was completed at approximately 90 days following the initial request. An additional 52 youth had discharged at Time 2, and 31 at Time 3. For these cases, providers were asked to complete a follow-up SDQ based on how the youth was doing just prior to leaving the program. An attempt was made to follow up with the case workers of youth who had discharged using the same strategy described above. To bolster case worker response, a member of the research team attempted to contact them by phone. However, due to COVID-19, many case workers were working outside of the office. Attempts to obtain follow-up from case workers were largely unsuccessful with only one case worker following through with a request for the youth to complete the SDQ.

MEASURES

The Strengths and Difficulties Questionnaire (SDQ), a 25-item behavioral screener, was used to measure changes in youth behavioral and emotional symptoms. The SDQ includes five subscales measuring emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior in the past 30 days. It includes a youth self-report and another version to be completed by adults such as parents, teachers, clinicians, or other professionals working with the youth being assessed. An example item from the youth form includes, "I try to be nice to people. I care about their feelings." The version of this item on the adult form is "[the youth] is considerate of other people's feelings." Items are rated on a 3-point scale where 0 = Not True, 1 = Somewhat True, 2 = Certainly True. QSA data from 2020 - 2021 were also used for this study to examine associations with SDQ scores.

DESCRIPTION OF DATA COLLECTED

Note that for each youth, some had an SDQ that was completed by the youth only, staff only, or, in some cases, a youth had a SDQ completed by the youth and another that was completed by a staff member who rated the youth's behavioral and emotional symptom severity. Table 2 displays the number of youths with an SDQ completed at each time point. A total of 118 youth had an SDQ completed at Time 1 representing an initial response rate of 24.9% (118/473). Rates of attrition from Time 1 to Time 2 were approximately 58.5% (69/118). There was a slight increase in SDQs completed from Time 2 to Time 3. Rates of attrition from Time 1 to Time 3 were 59.3% (70/118).

Table 2. Strengths and Difficulties Questionnaires Completed

Data Collection Period	Youth Only	RP Staff Only	Both Youth & Staff	Total
Time 1	48 (40.7%)	23 (19.5%)	47 (39.8%)	118 (100%)
Time 2	17 (24.6%)	27 (39.1%)	25 (36.2%)	69 (100%)
Time 3*	11 (15.7%)	34 (48.6%)	25 (35.7%)	70 (100%)

**Note – Time 2 RP Staff value includes 7 SDQs that were completed for youth with an SDQ at Time 1 who were missing an SDQ at Time 2.*

RESULTS

The analytic sample included 60 youth with an initial and, at least one follow-up SDQ. The sample included youth ages 11-17 and, as shown in Table 3, the majority identified as white and females. A

slightly larger percentage were in residential programs located in the Suncoast region, followed by Central, NE, SE, and NW, respectively. There were no youth from the Southern region in the sample.

Table 3. Sample Characteristics (*n* = 60)

Variable	Mean (SD)	N (%)
Age	14.97 (1.51)	
Race/Ethnicity		
Black		22 (37.9%)
White		36 (62.1%)
Gender		
Female		36 (62.1%)
Male		22 (37.9%)
Region		
Central		15 (25.9%)
Northeast		13 (22.4%)
Northwest		2 (3.4%)
Southeast		6 (10.3%)
Suncoast		22 (37.9%)

Note. Missing Gender = 2, Race/Ethnicity = 2, Region = 2.

Quality Standards and Youth Outcomes

To respond to the research question (Which standards are most strongly associated with youth outcomes?), paired samples *t*-test were performed to identify areas of significant improvement at each time point. Specifically, paired samples *t*-tests were run to examine changes from Times 1-2, Times 2-3, and Times 1-3 on the SDQ Internalizing subscale, Conduct Problems subscale, and Total Difficulties. Parametric bootstrapping was applied to correct for potential issues with non-normality and non-independent observations outside of expected autocorrelation within matched pairs. Table 4 shows mean differences and results for the youth self-report SDQ. Decreases in youth reported Internalizing scores were statistically significant when comparing Time 2 and Time 3 and Time 1 and Time 3. Youth self-reported conduct problems showed a small but statistically significant decrease from Time 1 to Time 2. Total Difficulties scores demonstrated statistically significant improvements when comparing changes across all three time points. In sum, the results showed that youth experienced significant improvements in self-reported internalizing problems, conduct problems, and total difficulties.

Table 4. Strengths and Difficulties Questionnaire – Youth Form

SDQ Scale	Time 1 – Time 2 (*n* = 42)			Time 2 – Time 3 (*n* = 32)			Time 1 – Time 3 (*n* = 35)		
	$M\ (SD)_{T1}$	$M\ (SD)_{T2}$	*p*	$M\ (SD)_{T2}$	$M\ (SD)_{T3}$	*p*	$M\ (SD)_{T1}$	$M\ (SD)_{T3}$	*p*
Internalizing	7.24 (3.38)	6.31 (3.99)	.07	7.42 (3.51)	2.82 (2.98)	.000***	7.56 (3.22)	3.11 (3.11)	.000***
Conduct Problems	2.21 (1.99)	2.74 (2.49)	.04*	1.90 (1.91)	1.60 (1.17)	.66	1.85 (1.77)	2.08 (1.75)	.77
Total Difficulties	22.05 (6.32)	19.07 (9.50)	.02*	23.09 (5.47)	7.53 (9.83)	.000***	23.43 (5.53)	9.03 (10.67)	.000***

Note. M = Mean, SD = Standard deviation, p = probability value.
**p* < .05, ***p* < .01, ****p* < .001

Table 5 shows mean differences and results of paired *t*-test for the provider completed SDQ. Provider ratings at Time 1 were low and indicated youth were admitting with few problems. This may be due to a tendency for youth to display fewer behaviors following initially being admitted. However, scores at Time 2, indicated elevated Internalizing problems and total difficulties in a range that might be expected of youth receiving care in residential settings. Provider ratings of youth internalizing behavior problems and total difficulties showed statistically significant improvements from Time 2 to Time 3. Additionally, although conduct problem scores remained low at all time points, means scores demonstrated statistically significant improvements from Time 2 to Time 3. Staff ratings were largely consistent with youth ratings demonstrating areas of significant improvement in internalizing problems, conduct problems, and total difficulties.

Table 5. Strengths and Difficulties Questionnaire – Provider Form

SDQ Scale	Time 1 – Time 2 (*n* = 32)			Time 2 – Time 3 (*n* = 52)			Time 1 – Time 3 (*n* = 36)		
	$M\ (SD)_{T1}$	$M\ (SD)_{T2}$	*p*	$M\ (SD)_{T2}$	$M\ (SD)_{T3}$	*p*	$M\ (SD)_{T1}$	$M\ (SD)_{T3}$	*p*
Internalizing	4.94 (1.98)	7.19 (3.89)	.000***	7.33 (3.98)	4.46 (4.15)	.000***	4.72 (2.01)	4.39 (4.06)	.62
Conduct Problems	2.94 (2.92)	2.41 (1.19)	.23	3.50 (2.82)	2.32 (2.51)	.003**	3.06 (2.48)	2.06 (1.26)	.04*
Total Difficulties	12.91 (2.80)	20.63 (5.57)	.000***	21.54 (6.28)	10.73 (11.51)	.000***	12.69 (2.74)	11.33 (10.89)	.45

Note. M = Mean, SD = Standard deviation, p = probability value.
**p* < .05, ***p* < .01, ****p* < .001

Next, a series of bivariate regressions using bootstrapped samples were performed to examine associations between youth reported symptom changes (i.e., total difficulties, internalizing problems, conduct problems) and the quality standards (item-level analyses; results not shown due to the large number of tests). The analyses identified several statistically significant associations between scores on the quality standards items and changes in youth symptoms on all three SDQ scales. This narrowed the list of standards associated with outcomes, which were then flagged for inclusion in subsequent analyses with added controls. To further examine which standards were associated with positive outcomes a series of stepwise regressions were performed, again applying parametric bootstrapping and controlling for SDQ scores. These analyses were performed separately for youth and provider completed SDQs and for each scale. For analyses of youth completed SDQs, all models controlled for SDQ scores at Time 1. For analyses of provider completed SDQs, all models controlled for SDQ scores at Time 2 as these were considered a more accurate reflection of youth symptoms.

Table 6 displays results for youth Total Difficulties. In total, six standards were statistically associated with lower total difficulties scores at follow-up. The first column shows that standard and specific item that was associated with the total difficulties scores and the last column identifies the QSA form. For example, from row one, higher scores on the item, "*My service plan includes goals that help prepare me to live with my family or on my own"* from the youth QSA, were associated with a 1.62 standard deviation decrease in youth self-reported total difficulties controlling for total difficulties at Time 1. This item measures the standard *Focus on safety, permanency, and wellbein*g from Domain 1. *Assessment, Admission, and Service Planning*.

Table 6. Associations between the Quality Standards and Total Difficulties - Youth

Standard/Item	Beta	p-value	QSA Form
D1.S. Focus on safety, permanency, and wellbeing. *My service plan includes goals that help prepare me to live with family or on my own.*	-1.62	.03**	Youth
D6.S. Respect for youth's privacy *I have privacy in my bedroom and bathroom.*	-0.22	.004***	Youth
D6.S. Trauma-informed approach *Staff or my therapist talk to me about trauma I have been through and how to deal with it better.*	-1.31	.07*	Youth
D6.S. Quality improvement approach *The program uses continuous quality improvement (CQI) in an on-going effort to evaluate and improve services.*	-0.44	.10*	DCW
D6.S. Actively monitor youth and milieu *Staff provide on-going supervision of youth according to program policies.*	-0.49	.02**	Director
D7.S. Measure and follow symptom reduction *The program focuses on reducing behavioral issues and symptoms in youth.*	-0.67	.03**	Director

Note. All models adjusted for SDQ Total Difficulties at Time 1; p-values obtained by parametric bootstrapping;
**p < .10, **p < .05, ***p < .01*

Table 7 shows that four standards were associated with lower provider ratings of youth total difficulties at follow-up. In these analyses, higher ratings of the youth QSA item measuring program performance on grievance processes demonstrated the strongest association with lower total difficulties.

Table 7. Associations between the Quality Standards and Total Difficulties - Provider

Standard/Item	Beta	p-value	QSA Form
D2.S. Effective crisis management *Staff effectively de-escalate crisis and behavioral incidents according to training and program policies, including documenting serious incidents.*	-0.47	.07*	Director
D3.S. Grievance process *Staff respond when I talk to them about things that I feel concerned about.*	-1.62	.002***	Youth
D6.S. Psychotropic medications are appropriately monitored *Psychiatrists monitor youth's psychotropic medication regimens at least once a month.*	-0.38	.01**	Director
D7.S. Provide vocational training opportunities *I was told by staff that I can receive job training for things like welding or cooking or other types of jobs if I want it.*	-0.22	.10*	Youth

Note. All models adjusted for SDQ Total Difficulties at Time 1; p-values obtained by parametric bootstrapping;
$*p < .10$, $**p < .05$, $***p < .01$

Table 8 shows that six standards were associated with lower youth reported internalizing problems at follow-up. In these analyses, higher ratings of the Director QSA item measuring program performance on the implementing culturally relevant and sensitive services demonstrated the strongest association with lower internalizing problems.

Table 8. Associations between the Quality Standards and Internalizing Problems - Youth

Standard/Item	Beta	p-value	QSA Form
D1.S. Respectful admission process *When I first got here, someone from this group home asked how I felt about coming here.*	-0.73	.07*	Youth
D1.S. Inclusive admission process *As much as possible, families are involved in creating service plans.*	-1.05	.02**	Licensing
***D4.S.* Culturally relevant and sensitive services** *The program supports youths' connection with their culture of origin.*	-2.26	.05**	Director
D6.S. Quality improvement approach *The program uses continuous quality improvement (CQI) in an on-going effort to evaluate and improve services.*	-0.73	.01***	DCW
D6.S. Psychotropic medications are appropriately monitored *Staff are aware of any adjustments to youths' medication, closely monitor dosage and any side effects, and report any concerns.*	-0.35	.001***	Director
D6.S. Actively monitor youth and milieu *Staff provide on-going supervision of youth according to program policies.*	-0.39	.04**	Director

Note. All models adjusted for SDQ Total Difficulties at Time 1; p-values obtained by parametric bootstrapping;
$*p < .10$, $**p < .05$, $***p < .01$

Table 9 shows four standards were associated with lower provider ratings of youth internalizing problems at follow-up. In these analyses, higher ratings by Lead Agency personnel of the QSA item

measuring program performance on the implementing assessment-driven services demonstrated the strongest association with lower internalizing problems.

Table 9. Associations between the Quality Standards and Internalizing Problems - Provider

Standard/Item	Beta	p-value	QSA Form
D1.S. Assessment-driven services *Evidence-based assessments are used to inform service planning.*	-7.44	.08*	Lead Agency
D2.S. Limited seclusion and restraint *Physical restraints and seclusion are used only in emergencies involving imminent safety risks.*	-0.47	.05*	Licensing
D6.S. Regular staff meetings to coordinate care *Regular staff meeting documentation includes attention to youth's progress, teamwork, and addressing program issues.*	-0.48	.02**	Director
D7.S. Measure and follow symptom reduction *Being in this program is helping me to feel and behave better than I did before I came here.*	-1.18	.03**	Youth

Note. All models adjusted for SDQ Total Difficulties at Time 1; p-values obtained by parametric bootstrapping;
**p < .10, **p < .05, ***p < .01.*

Table 10 shows one standard was associated with lower youth reported conduct problems at follow-up. There were no statistically significant associations between quality standards and provider ratings of youth conduct problems at follow-up.

Table 10. Associations between the Quality Standards and Conduct Problems - Youth

Standard/Item	Beta	p-value	QSA Form
***D8.S.* Connect family to community resources to support reunification** *Staff are helping me and/or my family find other programs and services we need to help me be successful after I leave here.*	-0.47	.03*	Youth

Note. All models adjusted for SDQ Total Difficulties at Time 1; p-values obtained by parametric bootstrapping;
**p < .10, **p < .05, ***p < .01*

In summary, the results comparing changes in youth behavioral and emotional symptoms, identified marked improvements in youth and provider reported internalizing problems and total difficulties that were both clinically significant and statistically significant. The results of these exploratory analyses found 17 standards were associated with lower internalizing problems, conduct problems, and total difficulties at follow-up. That is, youth placed in programs with higher performance ratings on these standards experienced significantly greater improvements. Standards from seven out of eight domains were identified as significant. Most of the significant standards were from Doman 6. *Program Elements* (six standards) followed by Domain 1. *Assessment, Admission, & Service Planning* (four standards), Domain 7. *Education, Skills, & Positive Outcomes* (two standards) and Domain 2. *Safe, Positive Living Environment* (two standards). Three standards were associated with multiple outcomes – D6. *Quality Improvement Approach* (youth Total Difficulties, youth Internalizing Problems), D6. *Actively Monitor Youth and Milieu* (youth Total Difficulties, youth Internalizing Problems) and D6. *Psychotropic Medications Appropriately Monitored* (provider Total Difficulties, youth Internalizing Problems).

Findings from this exploratory pilot provide preliminary evidence adding support that quality practice standards are associated with improvement in youths' emotional and behavioral symptoms. Specifically, higher scores on certain standards may be associated with improvements in youth internalizing problems, conduct problems, and total difficulties. Importantly, it should be noted that these analyses specifically aimed to respond to the research question, "Which standards are most strongly associated with youth outcomes?" Therefore, just because a standard was not reported as statistically significant in these analyses, does not mean that the standard (or practice it measures) is not related to youth outcomes. Further, the lack of associations with conduct problems may have been due to low conduct problem scores at all three time points. Study findings should consider the limitations. First, the study sample was small. Smaller sample sizes yield more error, increasing risk for failing to detect associations. Second, the study used a nonrandom design, thus, the results may not generalize beyond this study. Third, the follow-up period was limited to 60 and 90 days. Some youth in the sample had discharged while others remained in care. Extending the follow-up period may yield additional findings. Finally, sample attrition is a common challenge among studies using repeated measures designs. Attrition and obtaining a sufficient sample size was a challenge for this pilot study. Future analyses aimed at collecting youth outcomes data will require careful consideration of strategies to mitigate attrition and increase participation, in addition to utilizing a more rigorous study design to produce more robust findings. However, this study yielded findings to suggest there is value in further research identifying quality practices and how they impact youth outcomes in residential care.

Description of Quality Standards Assessment and Statewide Accountability System

Implementation of the QSA is led by the Department's regional licensing teams as part of the annual relicensure of residential group homes and shelters. The final version of the QSA is valid for evaluating core practices in all models of residential group care. It was also adapted and is valid for use with youth shelters.[3]

The QSA is completed online using Qualtrics Survey Platform. Studies have shown it provides reliable and valid results when completed by licensing specialists, Lead Agency case managers and placement specialists, residential care staff and directors, and youth. The childcare professional version includes all eight subscales with 96 items in total. Licensing specialist also complete a brief Trauma-Informed Care Checklist and Evidence-Informed Model of Care Checklist to guide their assessment of programs performance on related standards. The youth version includes seven subscales (Domain 5 was dropped due to poor performance during the pilots). Both the childcare professional and youth versions rate standards using a five-point scale. Qualtrics reports are generated to summarize assessment results that can be shared with the individual provider to support quality improvement. These reports are also generated to show overall performance across the state and by region. Finally, the QSA is being used as the core measures of Florida's SAS.

Conclusions

[3] *Items (Standards) with less applicability to shelters (e.g., youth educational progress) are excluded from the QSA when evaluating youth shelters.*

The Quality Standards and Quality Standards Assessment were created in collaboration with child welfare stakeholders throughout the state of Florida with the aim of improving the quality of residential care above and beyond licensing and accreditation standards. Development began in late 2015 and through a series of implementation pilots and studies, the QSA is fully developed, validated, and implemented statewide.

The outcomes pilot was an initial effort to systematically measure change over time in youth in residential care using a validated measure of youth behavioral and emotional symptoms. Despite challenges with a small sample and attrition, the analyses yielded preliminary results suggesting higher quality ratings on select standards spanning most domains were associated with improvement in youth behavioral and emotional difficulties.

Recommendations for Sustainability

The Department will continue to work with Dr. Boel-Studt and the FICW during a two-year transition period. During this time, Dr. Boel-Studt will provide an updated live webinar training with a component focusing on generating and interpreting QSA provider reports. A prerecorded training will be available for new hires and booster training purposes. Additionally, Dr. Boel-Studt will provide ongoing technical assistance to support the licensing teams in completing the QSA, build provider reports and export reports to a project landing page where they can be accessed by Department staff and licensing specialists. Dr. Boel-Studt will work with the Department data quality team on further report building (e.g., Dashboard) and any transition plans. Finally, trend data on programs' performance on the quality standards will be collected with baseline data collection beginning July 1, 2022. Recommendations for promoting maximum benefits of the QSA as part of the SAS, include focusing on defining and developing processes to systematically measure youth outcomes and to utilize QSA results to offer resources and training to support quality improvement efforts in Florida's residential group homes.

Reference

Boel-Studt, S. (2019). An assessment of quality standards for Florida's Department of Children and Families' licensed residential group homes: Results from the statewide pilot. Retried from https://issuu.com/fsuchildwelfare/docs/fr_an_assessment_of_quality_standards_for_florida_?fr=sODY3ODg1MzcwMw

Boel-Studt, S., & Huang, H. (2018). The development and validation of an assessment of quality standards for residential group care. Annual Report for 2017-2018. Retrieved from https://ficw.fsu.edu/sites/g/files/upcbnu1106/files/pdf-files/FR%20Development%20and%20Validation%20of%20the%20Florida%20Group%20Care%20%20Quality%20Standards%20Assessment-Fiscal%20Year%202017-2018.pdf

Boel-Studt, 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, 34,* 260-281.

Boel-Studt, S., Huang, H., & Dowdy-Hazlett, T. (2021). An Assessment of Quality Standards for Florida's Department of Children and Families Licensed Residential Group Homes: Annual Report 2020-2021. Retried from https://issuu.com/fsuchildwelfare/docs/fr_an_assessment_of_quality_standards_fy20-21_2?fr=sOWI1ODM1Nzg2MzI

Group Care Quality Standards Workgroup (2015). Quality Standards for Group Care. Father Flanagan's Boy's Home. Retrieved from www.myflfamilies.com/service-programs/child-welfare/residential-group-care-licensing/docs/1507-255%20Quality%20Standards%20for%20Residential%20Group%20Care%20Booklet.pdf