Innovations in Regulatory Science and Program Quality Assessment

Richard Fiene PhD

Table of Contents

1. Scientific Method: Finding the Rules that Work
2. Potential Solution to the Child Care Trilemma Revisited: Finding the "Right Rules"—The Holy Grail of Early Care and Education
3. The Instrument Based Program Monitoring Information System and the Indicator Checklist for Child Care
4. Regulatory Compliance and Quality Programs: Constraints and Opportunities for Integration
5. Regulatory Compliance Monitoring Paradigms and the Relationship of Regulatory Compliance/Licensing with Program Quality: A Policy Commentary
6. A Treatise on the Theory of Regulatory Compliance
7. A Comparison of International Child Care and US Child Care Using the Child Care Aware – NACCRRA (National Association of Child Care Resource and Referral Agencies) Child Care Benchmarks
8. The Child Care and Early Education Heart Monitor: The Intersection of Structural Quality and Process Quality Using the Contact Hour Metric As The Foundation
9. Key Quality Indicators: A white paper to introduce the concepts and application of key quality indicators in human care regulatory environments
10. Development of a Regulatory Compliance Scale
11. The Uncertainty–Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies
12. Improving Child Care Quality Through an Infant Caregiver Mentoring Project
13. Child Care Health Consultation Improves Infant Toddler Care

Innovations in Regulatory Science and Program Quality Assessment

Introduction and Overview

The 13 papers outline a significant paradigm shift in regulatory science and human services monitoring, particularly within the Child Care and Early Education (CCEE) field. This evolution moves away from traditional, uniform compliance checklists toward sophisticated, data-driven, and integrated systems that assess both foundational safety and the nuanced quality of developmental experiences.

The core of this new approach lies in the crucial distinction between **Structural Quality** (the health and safety standards mandated by licensing, or the "do no harm" principle) and **Process Quality** (the dynamic, interpersonal interactions between staff and children that foster development, or the "do good" principle). While structural quality establishes a necessary foundation, process quality is identified as the "heart" of a program's effectiveness.

This evolution is enabled by several key innovations:

- **The Theory of Regulatory Compliance (TRC):** A foundational concept that justifies a move toward differential monitoring, focusing on "substantial compliance" with a predictive subset of rules rather than 100% compliance with all rules.
- **The Key Indicator Methodology (KIM):** A statistical method for identifying the most predictive indicators of overall compliance, risk, and quality, enabling more efficient and targeted inspections.
- **Integrated Program Monitoring Systems:** The culmination of this work is the development of integrated systems that merge structural and process quality assessment into a single platform. The principal example is the **Child Care and Early Education Heart Monitor (CCEEHM)**, an advanced system that utilizes:
    - The **Contact Hour (CH) Metric** to provide a dynamic and more accurate measure of structural compliance with staff-child ratios.
    - A set of 10 validated **Program Quality Indicators (PQI)** to assess process quality across domains like curriculum, family engagement, and staff-child interactions.

    - **Artificial Intelligence (AI)** to conduct the thousands of observations necessary for reliable scoring, thereby reducing human bias and cost.

Finally, the approach argues that for these methodologies to be effective, organizations must adopt a process-based management approach that integrates regulatory compliance and quality programs, breaking down information silos and eliminating redundancies to improve overall efficiency and effectiveness.

---

1. The Evolution of Program Monitoring: From Compliance to Quality

The approach to licensing and oversight in human services has undergone a significant transformation, moving through distinct stages from basic compliance checks to holistic quality assessment.

- **Uniform Monitoring:** The traditional model, prevalent through the 1980s, where all rules and regulations were reviewed during every inspection. This approach proved to be inefficient and insufficient for capturing the true quality of a program.
- **Differential Monitoring:** A more targeted model that emerged with the development of the Key Indicator Methodology. This approach uses risk assessment and a subset of predictive rules (key indicators) to focus inspection resources on areas of highest concern, enabling abbreviated inspections.
- **Integrated Monitoring:** The current frontier, which explicitly incorporates quality assessment into the regulatory framework. This model evaluates both compliance with foundational rules and the effectiveness of developmental practices, offering a comprehensive view of a program's performance. The CCEEHM is a prime example of an integrated monitoring system.

This evolution is underpinned by the **Theory of Regulatory Compliance (TRC)**, which serves as a unifying framework. A key insight from TRC is the value of **substantial compliance**. Research has repeatedly shown a "ceiling effect" in structural quality data, where moving from substantial compliance to 100% compliance yields diminishing returns in terms of its relationship with process quality. This finding supports a policy shift toward issuing licenses based on substantial compliance, allowing for more efficient and focused regulatory practices.

2. Differentiating and Measuring Structural vs. Process Quality

A central theme across the sources is the conceptual and statistical distinction between two fundamental dimensions of quality in CCEE programs.

Feature	Structural Quality	Process Quality
Definition	The foundational, objective standards that ensure child health and safety.	The nuanced, dynamic interactions between staff and children that support learning and development.
Core Principle	"Do no harm."	"Do good."
Examples	Staff-child ratios, group size, safety regulations, number of violations.	Emotional climate, quality of language exchange, teacher engagement, problem-solving opportunities.
Typical Tools	Licensing rules and regulations.	Environmental Rating Scales (ERS), Classroom Assessment Scoring System (CLASS).
Measurement Type	**Binary (Dichotomous):** A rule is either in compliance or out of compliance (Yes/No).	**Ordinal/Interval:** Performance is measured on a scale (e.g., 1-7) that captures gradations from low to high quality.
Data Distribution	**Non-linear and positively skewed,** with a pronounced **ceiling effect.** Most providers cluster at or near full compliance, making it difficult to differentiate among high-performers.	**Linear relationship with a normal distribution.** Scores are more evenly spread, allowing for better differentiation across the full quality spectrum.

While both are effective at identifying the lowest-performing programs where violations and poor interactions often co-occur, the different statistical properties explain why correlations between structural and process quality are often modest. They are complementary constructs, with structural quality providing the necessary foundation upon which process quality can be built.

3. Methodological Frameworks for Assessment and Improvement

Several innovative methodologies have been developed to make regulatory assessment more reliable, valid, and nuanced.

The Key Indicator Methodology (KIM)

Originally developed for child care licensing in 1985 by Fiene, KIM is a statistical method used to identify a small subset of rules or standards that are highly predictive of overall performance. It has been successfully applied to identify:

- **Key Compliance Indicators (KCIs):** Predict overall regulatory compliance.
- **Key Risk Indicators (KRIs):** Predict risk to individuals in care.
- **Key Performance Indicators (KPIs):** Used in systems like Head Start monitoring.
- **Key Quality Indicators (KQIs):** Predict overall process quality, derived from systems like accreditation, professional development, QRIS, and observational tools.

Two critical components of this methodology are the **weighting of rules** and the **dichotomization of data**, which help mitigate false positives and negatives in licensing decisions.

The Uncertainty–Certainty Matrix (UCM)

The UCM is a 2x2 matrix designed to analyze and improve the reliability and validity of regulatory decisions. It compares a decision made by an inspector (e.g., "In Compliance" vs. "Not In Compliance") against the actual state of compliance (reality). The goal is to maximize agreements (++ and −−) while eliminating or reducing disagreements, which represent errors:

- **False Positives (+−):** Deciding a program is in compliance when it is not.

- **False Negatives (−+):** Deciding a program is not in compliance when it is.

The UCM can be used to identify and correct for randomness in decision-making and for positive or negative bias in individual assessors, thereby increasing the certainty of regulatory actions.

The Regulatory Compliance Scale (RCS)

To address the limitations of binary measurement in structural quality, the RCS was proposed. This is an ordinal scale (often 1–7) that aligns with the structure of process quality tools like the ERS. By creating categorical scoring for compliance, the RCS allows structural quality to be analyzed on more equal statistical footing with process quality, providing a more effective comparative tool than simple violation counts.

4. The CCEEHM: An Integrated System for Child Care

The **Child Care and Early Education Heart Monitor (CCEEHM)** is a comprehensive, integrated program monitoring system that combines the assessment of structural and process quality onto a single platform. It leverages AI to make this holistic assessment feasible, cost-effective, and efficient.

Structural Component: The Contact Hour (CH) Metric

The CH metric is a dynamic and superior alternative to static adult-child ratio and group size measurements. It calculates the total exposure time and density of contact between adults and children.

- **Methodology:** It requires answers to six simple questions about facility opening/closing times, staff and child arrival/departure times, and total numbers of staff and children.
- **Calculation:** An algorithm uses this data to build a trapezoidal model of the day's attendance pattern and calculates a CH value.
- **Analysis:** This calculated value is compared against a conversion table of compliant CH values for a given number of children and required ratio. A CH value that exceeds the table's threshold indicates non-compliance.

Process Component: The Program Quality Indicators (PQI)

The PQI tool consists of 10 validated indicators drawn from decades of research across various quality initiatives (accreditation, QRIS, professional development, ERS). These indicators measure the "heart" of quality.

Summary of the 10 Program Quality Indicators (PQI)

Indicator	Domain	Description
PQI 1	Staffing	Percentage of ECE III educators (AA/BA level) on staff.
PQI 2	Environment	Presence of a stimulating, dynamic, child-centered environment where children can access materials independently.
PQI 3	Curriculum	Use of a developmentally appropriate, emergent curriculum informed by individual child assessments.
PQI 4	Family Engagement	Opportunities for staff and families to build relationships through respectful, two-way communication.
PQI 5	Family Engagement	Regular, formal mechanisms (e.g., conferences, reports) for sharing information on a child's progress.
PQI 6	Staff-Child Interaction	(Preschool) Educators actively use materials and conversation to encourage children to communicate.
PQI 7	Staff-Child Interaction	(Infant/Toddler) Staff initiate engaging, turn-taking conversations (verbal and nonverbal) with infants and toddlers.
PQI 8	Staff-Child Interaction	(Preschool) Educators use language during daily events and play to develop children's reasoning and problem-solving skills.

PQI 9	Staff-Child Interaction	Educators listen attentively with their full focus when children speak.
PQI 10	Staff-Child Interaction	Educators speak to children with a consistently warm, caring voice and body language.

The Role of Artificial Intelligence (AI) and the CCEEHM App

The CCEEHM framework relies heavily on technology:

- **AI-Powered Observation:** AI is used to analyze video footage from classrooms, enabling the thousands of observations required to reliably score the PQIs and build the trapezoidal CH model. This process is more consistent and less biased than relying solely on human observers.
- **The CCEEHM App:** A software application makes the system accessible. It includes a CH calculator and a PQI assessment tool that automates the complex scoring protocols, providing a final quality classification (Low, Mid-Low, High-Mid, High).

A detailed hypothetical case study in the source material demonstrates the CCEEHM's diagnostic power, showing how a "High Quality" program met all CH and PQI standards (Total PQI Score: 36), while a "Low Quality" program was non-compliant on its CH metric and failed across all PQIs (Total PQI Score: 8).

5. Organizational Integration and Implementation

For these advanced monitoring methodologies to succeed, organizations must adopt a management system that reflects the same integrated, process-based philosophy. A significant barrier to effectiveness is the common practice of maintaining separate, siloed systems for regulatory compliance and quality programs.

Five Common Drags on Management System Optimization:

1. **Passive Reaction:** Passively accepting externally generated compliance targets without considering their value to the organization.
2. **Binary Formats:** Developing internal policies and forms with a simple binary (Yes/No) format that stifles measurement and improvement.

3. **Information Silos:** Maintaining two separate management systems—one for regulatory compliance and another for quality—preventing synergy.
4. **Narrative Structures:** Using dense, text-based procedures instead of clear, graphical process flowcharts and maps.
5. **Departmental Responsibility:** Assigning responsibility to entire departments rather than to individual process-owners, which obscures accountability.

The proposed solution is an **integrated, process-based management system**. This approach treats compliance and quality as interrelated processes. It uses process flowcharts to map activities and assign ownership at the individual level.

Benefits of an Integrated System:

- **Organizational Efficiency:** Eliminates redundancy by using shared procedures for activities like auditing, purchasing, and corrective actions.
- **Resource Optimization:** Saves time spent in audits and meetings and may lead to overall cost reduction.
- **Enhanced Knowledge:** Increases team members' understanding of the entire system, breaking down silos.
- **Improved Measurement:** Facilitates the establishment of process baselines, making it easier to set metrics and measure continual improvement.

The U.S. FDA's 2022 proposed rule to align its Quality System Regulation (21 CFR 820) with the international process-based standard ISO 13485 is cited as a catalyst and real-world example of this trend toward integration.

6. Supporting Evidence from Field Research

The effectiveness of targeted, quality-focused intervention is supported by a classic randomly assigned crossover clinical trial involving Child Care Health Consultants (CCHCs) in Pennsylvania.

- **Study Design:** Child care centers were randomly assigned to an immediate intervention group or a delayed intervention (contrast) group. The intervention consisted of a CCHC working with the center for 12 months to improve health and safety practices.
- **Findings:** The immediate intervention group showed a statistically significant improvement in their evaluation scores after one year (average score rose from 212 to 254; $p < .0001$). The delayed group showed no significant change during that

same period but demonstrated significant improvement after they received their CCHC linkage a year later.

- **Conclusion:** The study demonstrates that a focused, consultant-led intervention can produce measurable and sustained improvements in quality, reinforcing the value of systems that identify areas for improvement and link them to technical assistance.

A reprint from

American Scientist

the magazine of Sigma Xi, The Scientific Research Honor Society

Finding the Rules that Work

An emerging paradigm promises to close the gap between regulatory compliance scores and the quality of childcare services.

Richard Fiene

An old fable recounts how a father and son, taking a donkey to market to sell it, encounter a string of critical villagers who each inform the pair they're "doing it wrong." Their efforts to please each subsequent critic end, absurdly and tragically, with them carrying the beast of burden themselves, ultimately causing its death.

Like the advice of those villagers, regulations are proffered in the name of safety and good practice. And, like that father and son, programs that try to follow every single rule to the letter may soon find themselves too weighed down to achieve (or perhaps even recall) what they set out to do. As the saying goes, "When you're up to your behind in alligators, it's hard to remember that you set out to drain the swamp."

In my four decades as a regulatory scientist studying childcare, I've seen this pattern play out time and again: In the lead-up to evaluations, staff at perfectly compliant programs spend so much time dotting i's and crossing t's that they have little left over for working with classrooms or teachers, whereas staff at slightly less compliant facilities, though equally careful about observing rules, fuss less with paperwork and work more with teachers on improving skills and curriculum.

Needless to say, developmentally appropriate curricula change kids' lives; boasting a perfect record does not. This observation neither dismisses the 200 to 400 rules and regulations set by respective U.S. states nor undermines the importance of complying with them, either as individual rules or in the aggregate. And full compliance does improve safety. But, as data gathered by my research team repeatedly demonstrates, a vague, uncomfortable gap separates full, costly regulatory compliance from program quality.

It is never about more or fewer rules; it is about which rules are really productive and which are not.

Moreover, early care and education providers often voice concerns that licensing inspectors inconsistently administer and apply particular rules. At issue, then, are not regulations' overall value per se, but rather the value of individual rules relative to fanatical box-checking. Given their limited resources, how can the early care and education fields get the most bang for their buck?

Such a discussion is long overdue. The unequal worth of many general licensing and quality standards, including those driven by a regulatory political bent rather than empirical evidence, produce markedly uneven developmental outcomes for kids. Today, an outcomes-based scientific reference frame is already influencing the human services industry (childcare, child welfare, and child and adult residential services), particularly in the early care and education fields (childcare centers and family childcare homes for children between infancy and 12 years old). The point of my team's approach, which I call the *theory of regulatory compliance*, is not to ask whether we need more or fewer rules, or more thorough or less thorough compliance, but rather to evaluate which rules truly prove effective.

Modernizing Measurement

Regulatory scientists use tools, standards, and methodologies to assess the safety, efficacy, and quality of programs under government regulation. Ideally, they help regulatory agencies achieve the best possible public health and safety outcomes.

The regulatory science field has a lot of ground to make up. At about 30 years old, it lags its subject matter by a good century (Pennsylvania passed the first orphanage licensing law in the United States almost 140 years ago). Human services licensing grew slowly prior to the late 1960s to early 1970s, when American President Lyndon B. Johnson began the Great Society initiatives such as Head Start, which kicked off the rapid multipli-

QUICK TAKE

Contrary to historical assumptions, the quality of childcare programs does not increase linearly as their compliance with rules and regulations approaches 100 percent.

All-or-nothing, one-size-fits-all approaches to compliance and licensing generate skewed data, raise risks of false negatives and false positives, and burden staff with bureaucratic tasks.

Substantial regulatory compliance is an alternative approach that emphasizes compliance with the most productive rules, preserves safety, and allows staff to concentrate more on children.

DGLimages/Shutterstock

Staff of fully compliant childcare programs say they spend too much time box-checking and not enough working with teachers, whereas staff at slightly less compliant facilities, though equally scrupulous, bother less with form-filling and spend more time in the classroom. An outcomes-based substantial regulatory compliance approach lets licensors strike that balance.

cation of childcare programs. Those decades also saw human services, especially childcare, begin transforming from cottage industries, with program monitoring and measurement conducted qualitatively via case notes and anecdotal records, to more rigid systems that entailed oversight, case reviews, and state agency inspections. In the 1970s, these systems, which often varied from state to state, gave way to improvements brought by the Federal Interagency Day Care Requirements.

The watershed moment for regulatory science as it pertains to children's programs came in the 1980s. The previous decade's major childcare expansion in the United States had created a backlog of licensing assessments, caused unmanageable monitoring delays, and laid bare the logistical limits of case studies. These factors, combined with advances in computing, led states to introduce an empirical, quantitative, and instrument-based approach, complete with sophisticated software systems designed by state agencies and private vendors to track regulatory compliance and quality assessment data. Empirical evidence not only moved regulatory science from qualitative to quantitative analysis, it also revealed surprising patterns.

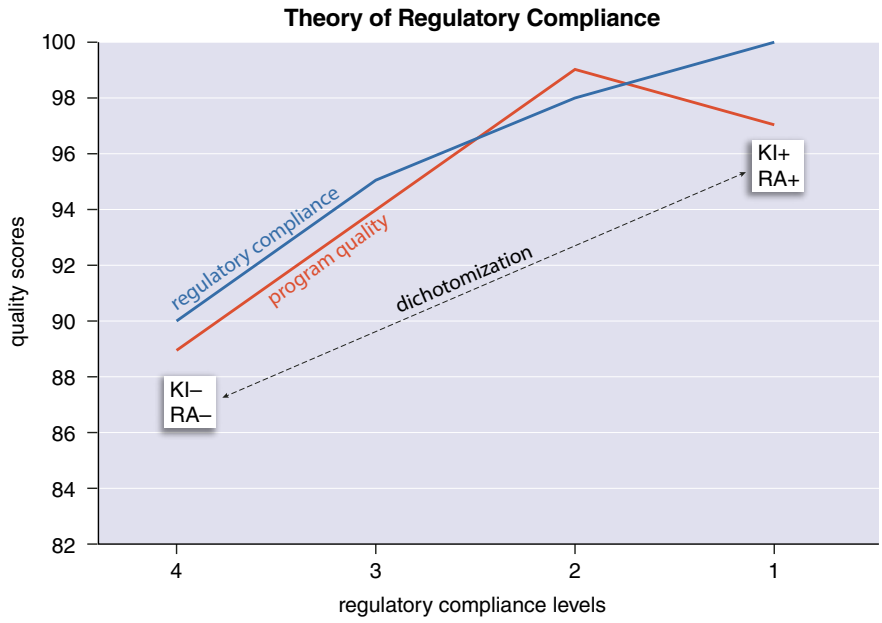
But first, some background: As the U.S. Department of Health, Education, and Welfare took over running the show for all U.S. early care and education programs in the 1970s, *uniform program monitoring* had become the rule. Uniform monitoring derived from the philosophical assumption that fuller regulatory compliance would produce, linearly, better quality across U.S. early care and education programs. As the former went up, so would the latter. From a public policy standpoint, this notion sounds aspirational, but sensible: Any licensing agency looks for service quality to increase as its rules, regulations, and standards are followed.

But as expert opinion and anecdotal evidence gave way to better-designed studies and empirical data, and as larger studies became possible thanks to data computerization by state licensing agencies, cracks appeared. When researchers compared violations found during licensing reviews and inspections to the quality of the violating programs, they found that a linear relationship did indeed exist between quality and compliance—but only as one moved from low compliance levels to *substantial regulatory compliance* (that is, 98–99 percent). Between that and 100 percent compliance, quality consistently plateaued and, as some 2010s replication studies suggested, even showed diminishing returns.

A New Paradigm

These results called into question the notion that state agencies should issue licenses solely to fully compliant programs. If, as data suggested, substantially compliant programs provided the same or better care as fully compliant ones, then clearly, we needed to rethink our program evaluation strategies.

In the United States, state licensing and regulatory agencies establish childcare regulations, but federal agencies such as the Office of Child Care and the Administration for Children



adapted from Richard Fiene

This graph shows the quality scores (*y*-axis) associated with four categories of regulatory compliance (*x*-axis, defined by the number of rules violations, ranging from 0 [Level 1] to 10 or more [Level 4]). Note that compliance scores (*blue line*) and quality scores (*red line*) rise together, but only until substantial compliance (99-97 percent compliance with all rules (Level 2) is reached. This finding argues for the adoption of substantial compliance as a standard, and for utilizing differential/relative monitoring to better capture nuances of quality and more efficiently allocate resources. The alternative—a punitive, gatekeeping licensing approach requiring full compliance (a yes/no proposition)—has led to highly skewed data. Here, the author has split (dichotomized) these skewed data into two extremes: Programs with regulatory compliance scores in the top 5-10 percent (*upper right*, labeled KI+/RA+ to indicate positive key indicator and risk assessment findings) and the bottom 5–10 percent (*lower left*, labeled KI-/RA-). The graph shows how scores in key indicators and risk assessment effectively predict program quality.

and Families also influence rules, as does Congress through its funding purse strings. Sometimes cities and counties, too, set regulations or standards, especially concerning physical environment, health, safety, and zoning. (Here, the term "regulations" means those defined by the National Association for Regulatory Administration's Licensing Curriculum.)

For an individual program or facility to operate, a state licensing agency must judge that it follows these standards. Examples include certifications for teacher qualifications, first aid, CPR, and the facility environment, along with requirements for ongoing training and professional development. State licensing staff evaluate compliance via inspections, document reviews, audits, and interviews, usually on a yearly basis. Inspections check for health, safety, cleanliness, educational standards, and staff-to-child ratios, as well as less obvious standards such as playground and transportation safety. Noncompliant programs may face fines, mandated corrective actions, training, or technical assistance, or may undergo license suspension or even permanent closure.

Licensing requirements vary depending on the childcare offered (such as family childcare homes, center-based care, or school-based programs), with larger centers typically facing more stringent requirements. Along with compliance ratings and violations issued by licensing inspectors, these facilities voluntarily seek ratings from quality initiative offices within human services agencies.

Here, and in my research, I primarily deal with center-based care programs, but the findings apply to other service types as well, such as family childcare homes and school-age programs, as well as human services categories such as child residential, child foster care, adult residential, and adult personal care homes. My data and research concern the relationship between quality and compliance, and how to improve it. They stem from studies of hundreds of programs I conducted at the state level from the 1970s through the 2010s, when I directed various research and training institutes at Pennsylvania State University. In these controlled and replicated studies, trained observers collected both regulatory data and program quality data from eight states, three Canadian provinces, and the U.S. Head Start program. The work ran the gamut, from site selection via stratified random samples, to dispatching data collectors to specific programs, to providing individual states with an overall blueprint describing how to conduct their studies.

Initially, the ceiling effect between regulatory compliance and program quality came as a surprise; we did not predict that full compliance would fail to outperform substantial compliance. It also drew pushback from the licensing field. Thus, I replicated the study many times over to assess my assumptions. But the finding persisted: Program quality scores rise with regulatory compliance until programs reach substantial compliance, after which quality declines. Although until 1980 states required childcare programs to show full compliance and zero violations, since 2015 most states have allowed licensing for facilities that are substantially compliant.

Differential Monitoring

If substantial compliance with some rules rather than full compliance with all rules best ensures the childcare program quality, then the question naturally arises: "Which rules?" Conceivably, some rules should weigh more heavily than others—say, the ones that data show most closely relate to safety and quality. Such is precisely the idea behind *differential monitoring*.

Differential monitoring emerged in 1979 during my discussions with federal agencies such as the Administration for Children, Youth and Families and the Children's Bureau, who felt dissatisfied with the traditional uniform monitoring approach. They knew about my team's work in Pennsylvania and invited me to give a series of talks to their staff. The result was a move away from the older, one-size-fits-all approach to differential methods focused on *key indicators* and *risk assessments*.

Key indicators are statistical predictors of overall compliance—rules that, if a facility follows them, strongly suggest they will follow other rules as

well. They very efficiently determine a facility's overall regulatory compliance without requiring a comprehensive inspection. Far from negligent, this approach works because not all rules are created and monitored equally.

Risk assessment focuses on those rules and regulations which, when breached, place children at greatest risk, such as rules that deal with supervision or hazardous materials handling, among others. Generally, jurisdictions, states, and provinces engage major early care and education stakeholders (service providers, parents, advocates, and licensing staff) in weighting rules or regulations based on their risks to children's health and safety. Commonly, participants assign weights via a *Likert scale*—a common survey and questionnaire tool that lets respondents indicate the strength of their agreement or disagreement (or, in this case, their assessment of risk) with a statement about attitudes, opinions, or perceptions. The weights range from 1 to 10, where 1 indicates little risk if a program fails to follow the specific rule or regulation and 10 corresponds to high risk. Rules heavily weighted as associated with sickness, injury, or death join the risk assessment rules measured by inspectors in every differential monitoring review.

As an aside, I should point out that full compliance remains the standard for maintaining health and safety. So why incorporate risk assessments into differential monitoring and, by extension, the substantial compliance paradigm, as its own separate metric? In truth, I had no such intention when I wrote my 1985 research papers about differential monitoring and the theory of regulatory compliance. Rather, risk assessment morphed from a way to provide the needed data variance for key indicator scoring into its own submethodology. As it found its way into the implementation of national standards and guidelines, risk assessment subsequently emerged as a separate methodology.

Our findings repeatedly show that using the combined methodologies of key indicator predictor rules and risk assessment rules to identify the "right rules" and to ensure compliance with them, rather than to seek full compliance, makes the differential monitoring approach the most effective and efficient program monitoring system. Also, studies show that abbreviated, targeted, and focused reviews take approximately 50 percent less time than comprehensive reviews.

Compliance Measurement Systems

scoring level	individual rule		aggregate rules	individual rule
scale	instrument based	*scale*	differential	integrated
7	full compliance	7	full compliance	exceeds compliance
–	–	5	substantial	full compliance
–	–	3	mediocre	substantial
1	out of compliance	1	low	mediocre/low

adapted from Richard Fiene

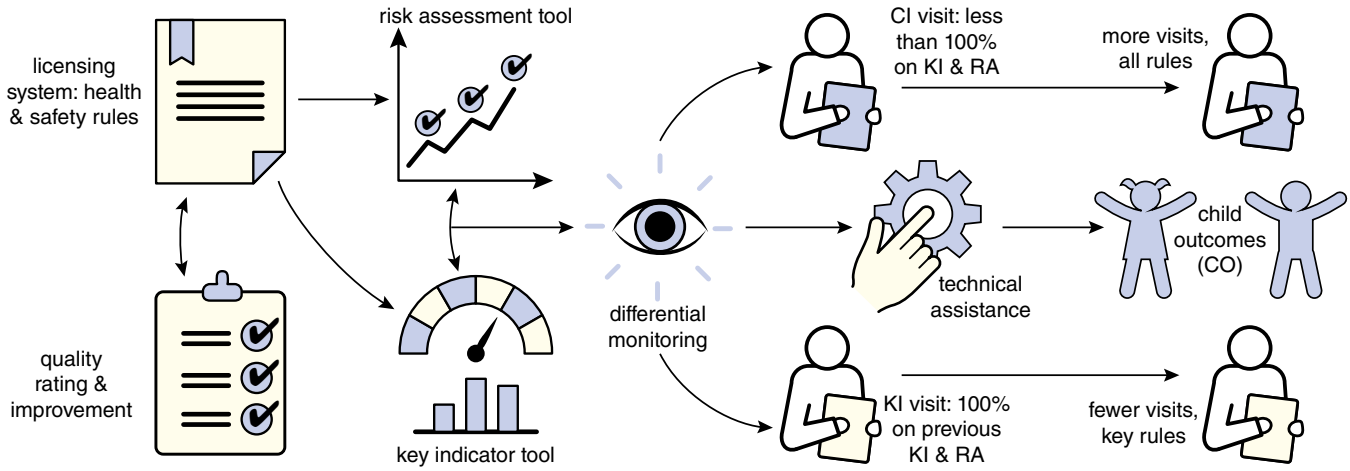
This table compares different approaches to measuring compliance: A licensing-focused approach in which programs are classified as either compliant or noncompliant based on rules violation counts, with no middle ground (*columns 1 and 2*), and a more nuanced ordinal approach using a Likert scale. This experimental metric, called the Regulatory Compliance Scale (*column 3*), is currently being tested at the aggregate rule level (*column 4*) and may be expanded to the level of individual rules (*column 5*) in the future. Note that aggregate rule scores are not equal to the sum of all individual rule scores because not all rules are created or administered equally.

Unfortunately, although many licensing bodies use risk assessment or key indicator methodologies, few use both. *Monitoring Practices Used in Child Care and Early Education Licensing*, a federal accounting of how states conduct program monitoring, reported that 10 states used key indicators, 17 states used risk assessments, and only one state used both. Hopefully, this pattern will change as the regulatory science field matures over the coming decades.

If, as data suggested, substantially compliant programs provided the same or better care as fully compliant ones, then clearly we needed to rethink our program evaluation strategies.

Since I first proposed it in the mid-1980s, the theory of regulatory compliance has faced numerous critics in the human services licensing field, especially among advocates of uniform monitoring and full compliance. Only after years of licensing validation studies conducted by my team and others repeatedly demonstrated that full compliance did not produce the highest quality did states begin licensing programs in substantial rather than full regulatory compliance. Today, although various U.S. states apply the differential monitoring review approach unevenly, nearly all have adopted the policy of granting licenses for substantial rather than full compliance. The latest revision of the legislation for the Child Care and Development Block Grant (a U.S. federal funding program that helps states, territories, and tribes assist low-income families access affordable childcare) cites differential monitoring as an alternative to uniform program monitoring.

Of all the approaches and methodologies that flow from the theory of regulatory compliance, differential monitoring most significantly alters the program monitoring, inspection, review, and licensing landscape. Its reviews occur just as often as do uniform monitoring assessments but focus specifically on rule breaches shown to place children at risk. That said, differential monitoring did not replace but rather supplemented its predecessor: Comprehensive reviews must still occur every three to four years to validate the performance of key indicators and risk assessment rules. But what does that report card look like in terms of analyzable data?



Barbara Aulicino

This illustration shows the various components that contribute to a differential monitoring approach and how agencies can use them to evaluate the effectiveness and validity of different approaches. Differential monitoring allocates resources based on risk assessment (client morbidity and/or mortality) and key indicators (rules whose compliance is strongly predictive of program quality). These data, provided by mandatory licensing processes and voluntary quality rating services, reveals which programs are highly compliant with key rules (though not all rules) and therefore require fewer visits versus programs that are less compliant and require additional visits and technical assistance to achieve similar child outcomes.

Rethinking Nominal Data

Traditionally, licensing data are categorical (sorted into groups such as "approved" or "denied"), unordered (there's no built-in way for such groups to be sequenced), and mutually exclusive (state agencies cannot simultaneously deem a facility both "approved" and "denied"). In statistical terms, such data are nominal, like a table listing cars by make or model; you cannot "do math" on such a table like you can on, say, on a table listing automobile curb weights and fuel economies. It is also binary: A program either follows a rule, or it doesn't.

Presently most jurisdictions deal in these absolutes and exclude gray areas. This approach, much like uniform program monitoring and full compliance, makes intuitive sense: We create rules and regulations because we believe in the value of following them, and because licenses mean nothing if licensees are not held to a standard. But here again, we must look deeper and ask, "What consequences follow from this either/or approach to measuring compliance, and who decides whether or not a particular box gets checked?"

Let's begin with the latter question. In an ideal world, judgments made by assessors would perfectly reflect a program's actual regulatory compliance state. But research that tests reliability and replicability in the licensing field empirically shows a concerning degree of disagreement when a second observer validates the decision regarding regulatory compliance. These disagreements suggest a worrying number of false positives and false negatives.

A false positive occurs when a program follows a rule or regulation, but the assessor rules that the facility is noncompliant (which might sound backwards, but the metric is *non*compliance, not compliance, so finding a false violation means finding a false positive). But even more concerning are false negatives, in which an evaluator says a program complies with a rule that it breaches, thereby placing clients at risk. Detecting false negatives is one of the chief reasons we periodically validate the predictive value of key indicator rules through comprehensive reviews.

As for the first question, the answer is simple: Nominal, binary licensing data is severely skewed. Upon reflection, the reason becomes obvious. When a regulated industry such as childcare mandates compliance before a program can operate and excludes gray areas, most facilities will achieve full compliance or lose their licenses. Because unlicensed providers don't last long, the childcare sector produces data that skew toward licensed programs. To grasp such skewed continuous or multicategory data, we must first dichotomize it into two distinct groups.

Such sorting into piles raises statisticians' hackles; unless carefully done, it accentuates differences and forces tradeoffs between precision and sensitivity, which can mean swapping false positives for false negatives. But the nature of licensing data—a skewed collection of mostly or fully compliant programs dumped in a single bucket—makes the split both necessary and warranted. By setting a threshold of certainty or agreement among evaluators, we can more effectively reduce false negatives, that is, cases in which evaluators say a program follows a rule when it doesn't.

This need becomes even clearer when one considers the demands posed by differential monitoring and its methodologies, key indicators, and risk assessments. For a program to receive licensure, it is not enough to ask if it "complies enough overall"; we must also know if it follows the specific rules that most ensure safety. By comparing highly compliant programs only with low-compliant programs, we accentuate the differences between the two and bolster our data analyses as well as overall safety. This comports well with licensing decision-making, which can consider a program compliant or noncompliant not only in aggregate, but with respect to *individual* rules.

Infusing Quality

The all-or-nothing approach to regulatory compliance and licensing fails as a standard because it generates skewed data, raises the risks of false negatives and false positives, and springs from a false assumption that program quality increases in step with 100 percent compliance. But I am far from the first

to notice that approach's weaknesses in evaluating how good a program or facility actually is. Indeed, its shortcomings helped drive the creation of a separate industry of voluntary accreditation programs such as the National Association for the Education of Young Children, state-run quality rating and improvement systems, and third-party tools and assessments. It's time we folded quality assessments into regulatory compliance.

I have already explained how the theory of regulatory compliance improves program quality and safety by focusing on substantial, not full, compliance and by using differential monitoring to ensure programs follow the most protective and impactful rules. But to further cast off the limitations and lopsidedness of a uniform monitoring and full compliance mindset, and to make room for data capable of tracking quality, we must also replace rigid either/or logic with a more nuanced ordinal measurement: a scaling technique.

Recall that assessors can evaluate compliance in two ways: They can consider aggregate rules—collections of rules that fall into categories such as staffing or safety practices—or individual rules. Each has its own studies and research literature. Research on aggregate rules from the 1970s, 1980s, and the 2010s established substantial compliance as a "sweet spot" of best outcomes and showed that the time had come to replace nominal metrics (such as "compliant" and "noncompliant") with ordinal ones (such as "98 percent compliant").

Inspired by this research, I have proposed replacing older nominal techniques with an ordinal scale like the Likert scale already used in quality measurements (usually but not always ranging from 1–7, with 1 being inadequate and 7 being excellent). This technique, currently under review by the National Association for Regulatory Administration, will help reviewers consider the importance of substantial compliance. Moreover, it will add the currently absent quality elements to each rule and regulation. However, this approach involves aggregate rules only; further research is needed to determine if the same shift from nominal to ordinal metrics should also occur at the individual rule level.

Should those findings bear out the value of evaluating individual rules via the 1–7 regulatory compliance scale, I propose that it should contain the following categories: exceeding full compliance, full compliance, substantial compliance, and mediocre compliance (*see figure on page 19*). These categories differ from the aggregate rule compliance scale currently under evaluation (full, substantial, mediocre, and low compliance) because aggregate compliance only considers health and safety elements, whereas an individual scale would also take quality into account.

Research supports the value of transitioning from uniform monitoring and full compliance to differential monitoring and substantial compliance. Practice has shown the value of retaining the older to help ensure the validity of the newer. Looking to the future, I believe we can further improve compliance evaluations by developing and evaluating *integrative monitoring*, which incorporates program quality into rule formulation and moves the key indicators from predicting compliance to forecasting quality.

The all-or-nothing approach fails as a standard because it generates skewed data, raises the risks of false negatives and false positives, and springs from the false assumption that program quality increases in step with 100 percent compliance.

Looking Forward

The regulatory compliance scale is a new and evolving metric. It transforms licensing data from a mere violation tally into a more useful and intuitive scale, one more consistent with the program quality measurements supported by research. Hereafter, I hope that the approach will incorporate quality measurements and more nuanced weighting into the evaluation of individual rule compliance. But discussions are just beginning, and this shift will pose a substantial challenge for agencies, which must also cope with the aftermath of the COVID-19 pandemic and a rising tendency toward deregulation.

The theory of regulatory compliance concerns the relationship between regulatory compliance and program quality, not health and safety, where full compliance remains the goal. It is, however, the preferred methodology for eliminating false negatives and decreasing false positives. Add to that the fact that the theory of regulatory compliance predicts a nonlinear relationship between compliance and quality but a linear relationship linking regulatory compliance and safety, and regulatory scientists clearly have our work cut out for us. Untying this knot will require greater collaboration between the historically siloed public policy worlds of licensing, accreditation, quality rating and improvement systems, and professional development systems.

I hope that the regulatory science field takes these paradigm shifts into consideration as it builds licensing decision-making systems and considers how states issue licenses. And although this work deals primarily with my own experience in the early care and education field, I wonder if other human service sectors, such as the foster care or child and adult residential areas, demonstrate similar patterns. Other disciplines that deal with regulations and compliance may similarly find it fruitful to discuss the nuances of their own evaluation metrics in order to achieve the best overall outcome with the most efficient use of limited resources.

Bibliography

Fiene, R. 2023. Licensing measurement and monitoring. National Association for Regulatory Administration licensing curriculum. Published online www.naralicensing.org/key-indicators

Fiene, R. 2022. Regulatory compliance monitoring paradigms and the relationship of regulatory compliance/licensing with program quality: A policy commentary. *Journal of Regulatory Science* 10:1–7.

Fiene, R. 2019. A treatise on regulatory compliance. *Journal of Regulatory Science* 7:1–3.

Richard Fiene is an emeritus professor of psychology at the Pennsylvania State University and has had a lengthy career in regulatory science and developmental psychology, in which his research has led to the development of national early care and education standards. He directs the Research Institute for Key Indicators Data Lab. Email: RFiene@RIKInstitute.com

Potential Solution to the Child Care Trilemma Revisited

Finding the "Right Rules"—The Holy Grail of Early Care and Education

by Richard Fiene

Rules and regulations: Can't live with them, Can't live without them.

How often have you heard this statement? I have heard it a great deal in an early care and education career that has seen six decades of discussion about what is the right mix of rules and regulations, the basic protections for children while in out of home child care. Recently, in the early care and education field, there has been a great deal of discussion about deregulation of early care and education standards/rules/regulations in order to have increased access to child care (National Association for the Education of Young Children: NAEYC, 2024). This discussion or controversy has been going on for a long time, it is nothing new. I remember back in the early 1970's when I was directing the Mary Elizabeth Keister Infant Toddler Demonstration Center at the University of North Carolina at Greensboro and there were discussions about the revision to the Federal Interagency Day Care Requirements (FIDCR). What was the right mix, the balance of protections and quality enhancements for young children in child care that the Federal Department of Health, Education and Welfare wanted to promulgate nationally.

Rick Fiene
Exchange Leadership Initiative
exchangepress.com/leadership

Rick Fiene has been working in prevention research for more than 50 years, focusing on child care policy, child care quality and human service regulatory administration and science. He initially started in prevention research because he wanted to have a positive impact on children's lives. He believed that researchers were not paying enough attention to the licensing of child care programs and decided to base his career on improving this area of research.

Fiene is regarded as a leading international researcher/scholar on human services licensing measurement and differential monitoring systems. His theory of regulatory compliance has altered human services regulatory science and licensing measurement dramatically in thinking about how best to monitor and assess licensing rules and regulations through targeted and abbreviated inspection methodologies: differential monitoring, risk assessment, and key indicators. He received the National Association for Regulatory Administration President's Award in 2015 and the Pennsylvania Association for the Education of Young Children's Distinguished Career Voice for Children Award in 2020.

But I think there is a better way to deal with this discussion which is driven by regulatory science and the empirical evidence that has emerged over the past 50 years. Let's take this discussion out of the political domain and place it where it needs to be, firmly within the newly emerging regulatory science field and focus on regulatory compliance. There is a theory of regulatory compliance (Fiene, 2019) getting kicked around a good deal in the human services regulatory science field that has upended the way we make licensing decisions. The theory has been empirically proven in several studies throughout the U.S. and Canada (Fiene, 2025). The theory simply states that substantial regulatory compliance with child care rules and regulations may be equivalent to full (100 percent) regulatory compliance with all child care rules and regulations. From a public policy and licensing decision making point of view, it changes program monitoring from a uniform one-size-fits-all approach to a more targeted and focused differential monitoring approach that looks at risk assessment and prediction of overall compliance with rules and regulations (Fiene, 2025).

So that is the theory but where do we start at a practitioner level? If we start at the baseline of early care and education

quality, then licensing and *Caring for Our Children (CFOC): The National Health and Safety Performance Standards, 4th Edition* (AAP, APHA, NRCHSCC, 2019), published by the American Academy of Pediatrics (AAP), the American Public Health Association (APHA), and the National Resource Center for Health and Safety in Child Care (NRCHSCC) is a good place to start because the CFOC is considered the default set of health and safety standards in the early care and education field. The standards were first published in the early 1990's and have been refined through several revisions and editions over the past several decades in response to the everchanging early care and education research literature related to health, safety and program quality. For over 30 years, the standards have been the reference for state child care licensing agencies as they think about promulgating new or revised rules/regulations/standards in their respective states. It is based upon the latest science in developmental psychology, pediatrics, and public health fields related to early care and education settings. The standards have been peer reviewed by expert technical panels representing all of the above areas of developmental psychology, pediatrics, public health, environmental health, etc. But it is a daunting document, over 700 standards are within this reference manual for the early childhood field.

Advocates point to *Caring for Our Children* (AAP, APHA, NRCHSCC, 2019) as the go-to-document because it provides a solid floor to quality while building on this base to demonstrate best practices. Others, mostly in the political arena, point to it as an example of over-regulation, too many rules to follow. But let's not forget what *Caring for Our Children* (AAP, APHA, NRCHSCC, 2019) is all about, protecting our children while in out of home care. Access to child care is important for many families, as is access to quality child care, as is access to safe and quality child care. Navigating these all is a delicate and challenging balance.

So, what is a potential solution to the child care trilemma? Let's look at regulatory science for potential guidance. As I said earlier, regulatory science is an emerging field, it is not well developed as the other physical and social sciences but it is making tremendous strides in the past 20-30 years. There are two parallel tracks, one dominated by the pharmaceutical industry and the other in the human services, in particular, in early care and education. In the pharmaceutical arena there is more concern about clinical trials and the efficacy of drugs and protection from side effects for individuals; in the human services arena there is more concern about protections from harm related to caregiving. And this is where regulatory science came into play with a new methodology in the human services that was emerging around risk assessment and key indicator rules/regulations to make monitoring more effective and efficient by focusing on risk to children and prediction of overall regulatory compliance (Fiene, 2019, 2025).

Initially there was more focus on the risk assessment methodology to determine if certain *Caring for Our Children* (AAP, APHA, NRCHSCC, 2019) standards placed children at increased risk of morbidity and mortality if regulatory non-compliance occurred. The resulting document, *Stepping Stones to Caring for Our Children* (NRCHSCC, 2019), came about based upon this risk assessment rule methodology. It took the over 700 standards to distill it down to approximately 120 standards. It became a much more manageable document that state licensing agencies could use as a guide in revising and promulgating rules and regulations.

Later in the development and evolution of *Stepping Stones to Caring for Our Children* (NRCHSCC, 2019), again borrowing from the regulatory science field, the key indicator rule methodology was utilized to determine if there were a smaller set of standards that had more of a predictive value in protecting children when it came to regulatory compliance in an overall sense. This resulted in *Caring for Our Children Basics* (CFOCB) (ACF, 2015) (approximately 65 standards) which was originally proposed as a voluntary set of standards for all early care and education. I think it was a good idea back when it was first proposed in 2015 and I still think it is a good idea. To many, 65 standards may still sound like too many standards but these standards form the basis for the quality and safety arm when it comes to the child care trilemma, and indirectly impact accessibility and affordability. The more standards to meet, the greater the cost for programs which can make it more difficult for parents to access available care. As quality increases, so does cost while accessibility decreases based upon what parents can afford.

Let's begin here in attempting to address a revised solution to the child care trilemma. In this discussion about where the child care field is headed and the most recent call for deregulation (Hechinger Report, 2024)(NAEYC, 2024), let's pivot and think about using *Caring for Our Children Basics* (ACYF, 2015) as our point of discussion rather than arbitrarily removing rules with this deregulation mind set because it is politically expedient. Let's be driven by the empirical evidence and the science which *Caring for Our Children Basics* (ACYF, 2015) is derived from solid regulatory compliance methodologies of risk assessment and key indicator rule/regulatory/standard identification. See how your state's child care rules size up with *Caring for Our Children Basics* (ACYF, 2015) in making sure that at the very least all these standards are in place. Templates from regulatory science have been developed

to do this comparison (Fiene, 2025). As a very important footnote regarding these standards, they were developed by a cross-representation of medical experts, early care and education experts, child developmental experts, public health and environmental experts. So all disciplines having an impact on child care services were well represented and consulted in the development of the standards.

Then once this is done in the aggregate, begin to look at the individual standards within *Caring for Our Children Basics* (ACYF, 2015). Let's be honest, probably the most discussed standard is staff-child ratios and group sizes. It has the greatest impact on cost (staff), numbers (children), and quality. This has been clearly demonstrated in the research literature for over 50 years. Nothing has changed, it was the focal point back in the 1970's (Abt, 1979) and it is today (Fiene & Stevens, 2021). But let's think outside the regulatory compliance box for a minute and maybe we do not look at staff-child ratios in isolation but cross it with another standard/rule such as the qualifications of staff and suggest an alternate rule where staff-child ratio can be increased slightly but only with the most highly qualified staff?! Like I said, let's think outside the regulatory compliance box. And while we are there, the fee that is attained by the program with the additional child should go to the more qualified staff as an add on to their salary. Yes, they have an additional child but they also have the revenue generated as a salary increase with the addition. This above approach I suggested in a Child Care Information Exchange article back in 1997 in how this approach could be utilized effectively as a potential solution to the child care trilemma (Fiene, 1997).

As with staff-child ratio and group size, we perform the same type of critical analysis utilizing the empirical regulatory compliance data available to make changes in the existing set of rules. As has been pointed out in the regulatory science research literature, regulatory compliance with rules is a measurement issue, so it should be solved in a corresponding way, use the data, do not ignore the empirical evidence and leave it up to the whims of the political process to determine what stays and what gets pitched. For the interested reader, there are several studies completed by the National Association for Regulatory Administration (NARA) which can guide one in determining how best to use data to make these decisions. These can be found at naralicensing.org/key-indicators

The point of this research abstract position paper is for us to take a step back and avoid a knee-jerk reaction to dealing with the child care crisis and that the only solution is to increase availability and affordability at the expense of health, safety and quality via deregulation (NAEYC, 2024)(Hechinger Report, 2024). We now have an emerging regulatory science (Fiene, 2025) to guide us and I hope we use it for making educated and informed choices as we move forward in attempting to solve the continuing child care trilemma.

References

Abt Associates. (1978). *Day Care Centers in the U.S.A National Profile 1976-1977*, Boston, Massachusetts.

American Academy of Pediatrics and the American Public Health Association and the National Resource Center for Health and Safety in Child Care: AAP, APHA, NRCHSCC. (2019). *Caring for Our Children: The National Health and Safety Performance Standards*. Itasca, Illinois.

Administration for Children and Families: ACF. (2015). *Caring for our Children Basics*. Washington, DC.

Fiene, R. (2025). Finding the Rules that Work: An emerging paradigm promised to close the gap between regulatory compliance scores and quality of childcare services. *American Scientist, 113* (January-February 2025).

Fiene, R. (2024). *Regulatory Compliance and Monitoring Systems*. Research Institute for Key Indicators and National Association for Regulatory Administration.

Fiene, R. (2019). A Treatise on the Theory of Regulatory Compliance. *Journal of Regulatory Science, 7*, 1-3.

Fiene, R. (1997). Searching for a Solution to the Child Care Trilemma. *Child Care Information Exchange*, September, 1997, 57-59, Redmond, Washington.

Fiene, R. & Stevens, S. (2021). *Contact Hours as a New Metric Replacing Group Size and Staff-Child Ratios as well as a New Metric for COVID19 Thresholds*. Research Institute for Key Indicators/Penn State.

Hechinger Report. (2024). *Investigating How Kids are Harmed by Lower Child Care Standards*. New York, New York.

National Resource Center for Health and Safety in Child Care: NRCHSCC. (2019). *Stepping Stones to Caring for our Children*. Boulder, Colorado.

National Association for the Education of Young Children: NAEYC. (2024). *The Costs of Deregulating Child Care*. Washington, DC.

The Instrument Based Program Monitoring Information System and the Indicator Checklist for Child Care

Richard Fiene
Office of Children Youth and Families
Commonwealth of Pennsylvania
Mark Nixon
Children's Services Monitoring Transfer Consortium, Washington, D.C.

ABSTRACT: The Instrument Based Program Monitoring Information System (IPM) and the Indicator Checklist (IC) are two tools for the state management of child day care services. A methodology for monitoring interviews and site visits to child day care programs is described. An integral feature of IPM is a system of assigning weights to the questions or items so that scores reflect the relative importance of state regulations. An Indicator Checklist is a questionnaire or checklist that contains selected, predictive items from a longer, comprehensive instrument that a state uses to monitor child day care providers' conformance to state day care regulations. An Indicator Checklist contains items that have been determined to be most effective in discriminating between providers that typically receive high overall scores on the comprehensive instrument and providers that typically receive low overall scores.

For nearly half a century, state governments have accepted responsibility for ensuring that those who care for children in their home and in day care centers meet minimum requirements for health and safety. During the past decade as the amount of state and federal funds for day care have grown, states have taken an active role in monitoring (1) the ways in which day care providers administer their programs, and (2) the quality of the services provided to children for whose care the state is paying.

Nationally, day care is big business. It is estimated that currently there are more than 118,000 licensed providers who serve an estimated 1.2 million children every day. The stakes in assuring that these children are well served are high, both in terms of public health and safety and from the viewpoint of enhancing the growth and development of America's most precious resource, its children. It is estimated that $6.3 billion dollars are spent annually on day care services.[1]

Reprints should be requested from Richard Fiene, Directory of Research and Information Systems, Office of Children, Youth, and Families, 1514 North Second Street, Harrisburg, PA 17102.

[1] Day care services include group day care centers serving 12 or more children, group day care homes serving 6–11 children, and family day care homes serving 5 or fewer children. Head Start & nursery school programs that operate for part day are included in day care services definition.

Child Care Quarterly, 14(3), Fall 1985

However, in monitoring these services, states spend less than one percent of their day care funds each year to ensure that providers comply with regulations or meet quality guidelines.

This article describes an approach in monitoring child day care services called: Instrument Based Program Monitoring (IPM). An IPM differs substantially from the more common approach to monitoring: narrative site visit reports used by most states. The narrative report approach usually includes a site visit to each provider and the preparation of a summary of observations and interpretive and evaluative comments about the monitor's findings. These reports are time consuming to prepare, and often difficult to summarize succinctly for policy makers and administrators. This article describes an alternative to the narrative site report.

Forces Changing the Regulatory Environment

The job of state agencies in program monitoring is currently changing in response to powerful forces in American society, especially at the level of state government.

First, there is the continuing need to assure parents that their children will not be subjected to unsafe day care environments and that day care providers who receive state funds are meeting the terms of their contracts with the state by providing quality services. Quality services are defined as day care services that promote sound child development principles and do not only ensure that children are in healthy and safe child care environments. Public accountability requires that the state entertain a dual purpose, one is to monitor compliance with state regulations; but secondly and equally important, there is a strong need for the state to ensure that quality child development services are supported and provided.

Gwen Morgan's (1980) work is particularly helpful in providing direction regarding the relationship between licensing and funding criteria. A Model presented by Morgan (1980) clearly delineates a regulatory continuum where day care licensing is considered as the floor to quality with accreditation as the standard of quality for which model day care programs strive. Recent efforts by the National Association for the Education of Young Children (Center Accreditation Project (1983)) and the Children's Services Monitoring Consortium (Child Development Program Evaluation Scale (1984)) have helped to support this move towards accreditation and the measurement of quality in early childhood programs. These efforts take on additional meaning given the direction from the federal government to pass as much of the responsibility for monitoring early childhood programs to the states.

Second, the fiscal cutbacks that are now occurring in many states will almost certainly increase the pressure on state agencies to operate as efficiently as possible. Cutbacks in staff across agencies are likely, even as workloads increase. These factors will force states to streamline their regulatory enforcement and monitoring efforts in all areas, including day care and children's services. A promising approach attempted in some states is moving from a licensing to a registration system. In a registration system, the locus of control for the regulatory process is shifted from the state to the provider level—the provider is responsible for assuring that s/he meets all registration requirements.

Third, the role of the state in regulating private sector organizations is changing. There are now active pressures to reduce the general level of state regulation with a view toward encouraging private market forces in the production and allocation of goods and services. Further, there is a commitment in a growing number of states to reduce the extent of the Federal Government's involvement, including federal funding and accompanying regulatory requirements, in several areas, notably human services (The moratorium placed on the Federal Interagency Day Care Requirements is a specific example which was supported by a number of states).

Fourth, many states are actively seeking ways to reduce the burden on the private sector of the compliance monitoring activities that are perfomed by the state. For those regulations that continue in force, many states will be examining approaches that simplify monitoring procedures and make them less onerous for providers. This is particularly true for day care services, which are often provided by individuals or organizations that may have little experience coping with regulations.

IPM as a Response to These Forces

One approach that states have used to cope with these forces is the development of Instrument-Based Program Monitoring Systems—(IPMs).

As the name implies, an IPM system incorporates three distinguishing characteristics: *First,* it is instrument-based. The system uses checklists or questionnaires that contain highly specific questions. These questions usually correspond directly to the state's regulations or other requirements (e.g., fiscal requirements). *Second,* it supports program monitoring. In its broadest sense, program monitoring is the management process of conducting periodic reviews

or inspections to ensure that certain activities, such as the provision of day care service, meet acceptable criteria, and the process of effecting corrective action where required. Program monitoring may include one or some conbination of:

1. Licensing reviews (Table 1 gives a listing of items taken from Pennsylvania's IPM at the licensing and minimal standards level);
2. Contract compliance reviews; and
3. Evaluations of program quality that go beyond minimum requirements to health and safety. A specific example that may be helpful is taken from the *California Child Development Program Quality Review* (1982) Instrument. What follows is a sampling of the Table of Contents:

PROGRAM QUALITY SUB SCALE

A. GOALS AND OBJECTIVES OF CHILD DEVELOPMENT PROGRAM ARE EVALUATED AT LEAST ANNUALLY BY THE STAFF AND PARENTS AND ARE MODIFIED AS NEEDED
B. TEACHING STAFF HIGHLIGHTS EACH CHILD BY SHARING INDIVIDUAL ETHNIC AND CULTURAL BACKGROUNDS—EMPHASIS IS PLACED ON CARE-GIVER OBSERVATIONS.
C. THE GOALS, OBJECTIVES, AND PROCEDURE FOR IDENTIFICATION OF CHILDREN'S NEEDS ARE EVALUATED AT LEAST ANNUALLY BY STAFF AND PARENTS (Fiene, 1984).

Third, IPM is a comprehensive system. It is part of a group of related steps such as on-site reviews, corrective action, follow-up reviews, and summarizing and reporting results that are used recurrently to accomplish the task of compliance monitoring. Program, fiscal, and statistical components can be linked quantitatively to constitute a comprehensive IPM system for day care. A new software decision support system (Watson, Fiene, & Woods, 1984) based on IPM is being developed for micro-computer technology and is being pilot tested in Michigan Department of Social Services, and Texas Department of Human Resources. When the IPM system is used in this linked fashion, it provides the basis for monitoring child day care Vendor & Voucher Delivery systems.

The advantages of an IPM system that are responsive to the changes mentioned earlier include: consistency, coverage of all regulatory areas, clear expectations simplified monitoring procedures,

TABLE 1

Pennsylvania Child Development Program Evaluation

Specific Items Within Identified General Areas

General Requirements	
1. Relevant approvals	4. Child abuse reporting procedures
2. Insurance coverage	5. Provision for special services
3. Parent participation	
Staffing Standards	
1. Qualifications of staff	staff requirements
2. Responsibilities	4. Staff health requirements
3. Adult/child ratio and minimum	
Employee Records	
1. Evidence of qualifications and references for staff	
Building & Site	
1. Appropriate indoor and outdoor square footage per child	materials
2. Characteristics of play areas	5. Cleanliness
3. Sanitary facilities	6. Screening of windows and doors
4. Storage of medicine and	7. Heating apparatus
	8. Educational materials available
Equipment	
1. Condition and placement of equipment	2. Swimming regulations
	3. Napping rules
Program for Children	
1. Evidence of written program plan with developmental activities	special needs children
2. Discipline	4. Sanitary habits developed
3. Identification and referral of	5. Infant/toddler stimulation
	6. School-age requirements
Food & Nutrition	
1. Menu requirements	3. Utensils
2. Infant formula rules	4. Special diet considerations
Transportation	
1. Vehicles all licensed and inspected	4. Restraint of children
2. Insurance coverage	5. First-aid kit materials
3. Adult/child ratio	
Child Health	
1. Requirements of health records	4. Medications
2. Emergency contact information	5. Procedure for ill children
3. Medical emergency procedures	6. First-aid requirements
Staff Health	
1. Procedures for staff illness	2. Physical requirements for infant caregivers
Procedures & Applications	
1. Pre-admission policy	3. Requirements of day care agreement
2. Requirements for child's application	
Child Records	
1. Frequency of updating records	4. Parental rights to records
2. Confidentiality	5. Procedure for release of information
3. Information to be included in child's records	6. Use of records after termination of service

and potential for cost efficiencies. With an IPM system, the same questionnaire or checklist is used with all providers, and there is less opportunity for individual bias in reporting results. Similarly, basing the questions or checklist items explicitly on the regulations or other requirements makes it possible to ensure that all areas are covered adequately. Having a clear set of questions that are known to both monitoring staff and providers reduces the possibility of misunderstandings and misinterpretations concerning the results of the review. Finally, standardized procedures for administering the questionnaire and processing the results can simplify the state's monitoring task and reduce the time, cost, and burden of monitoring both to the provider and to the state.

Four agencies (Pennsylvania's Office of Children Youth and Families, West Virginia's Office of Social Services, California's Office of Child Development, and New York City's Agency for Child Development) that are part of a consortium for improving the monitoring of children's services (Children's Services Monitoring Transfer Consortium) have experienced significant improvements in provider satisfaction with monitoring efforts and have, in some cases, achieved more efficient allocations of resources for day care and day monitoring. Pennsylvania has experienced substantial cost savings by linking the results of their IPM system to the state's fiscal and statistical information systems (See Figure 1). The state was able to set a ceiling on

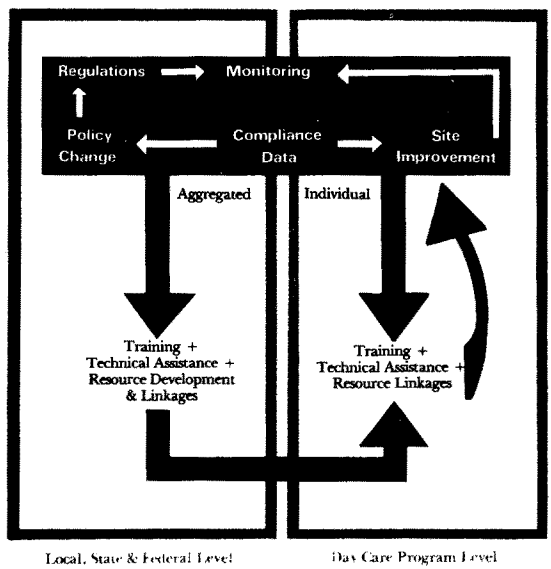


FIGURE 1

day care funding that did not jeopardize program quality, and used the funds that were formerly given to high-cost providers to improve services of other providers on a targeted basis. The state saved approximately $5 million in day care funds while maintaining the quality of day care services, and it did so without major resistance from the provider groups. California has been able with its IPM system to begin automation of its licensing and program quality instruments and linking these data with unit cost and service information on providers. In the development of the program quality instruments, a representative sample of providers from across the state played a critical role in the development and implementation of California's IPM system. These links are providing the basis for a child development, decision support system for the Office of Child Development in California.

Indicator Checklist Improves IPM Systems

Very recently, a number of states (Pennsylvania, West Virginia, Michigan, California, Texas, and New York) have begun experimenting with what has been called an "Indicator Checklist." Simply defined, an indicator checklist is a questionnaire or checklist that contains selected items or indicators from a longer, comprehensive instrument that is used as part of an IPM system. The items on the checklist are those that have been determined to be most effective in discriminating between providers that typically receive high overall scores on the comprehensive instrument or provide a high level of quality care and providers that typically receive low overall scores or provide low level of care (Figure 2).

Because of their value in distinguishing between providers who are in compliance and those that are out of compliance, the items on the in-

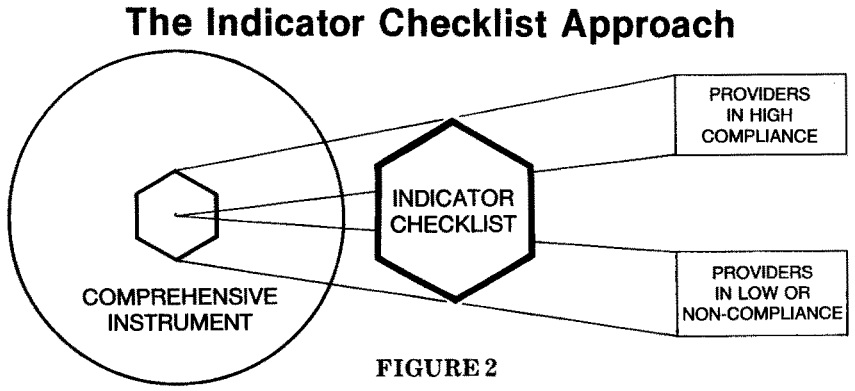


FIGURE 2

dicator checklist have been called "predictor" items. That is, they are a subset of items from the longer instrument that have a strong ability to "predict" the results that would have been obtained had the comprehensive instrument been administered to a given provider. In four of the states mentioned above, the average length of their respective Indicator Checklist's have been approximately 25 items. This compares with the average of approximately 200 items on their respective comprehensive instruments. The relationship between the scores obtained on the state's Indicator Checklists and their comprehensive instruments have been extremely high. When a Pearson's Product Correlation Coefficient was calculated on the Indicator Checklist and the comprehensive instrument for each state the correlation coefficients were always at a r=+.80 or higher (See Figure 2a for a graphic display of West Virginia's data).

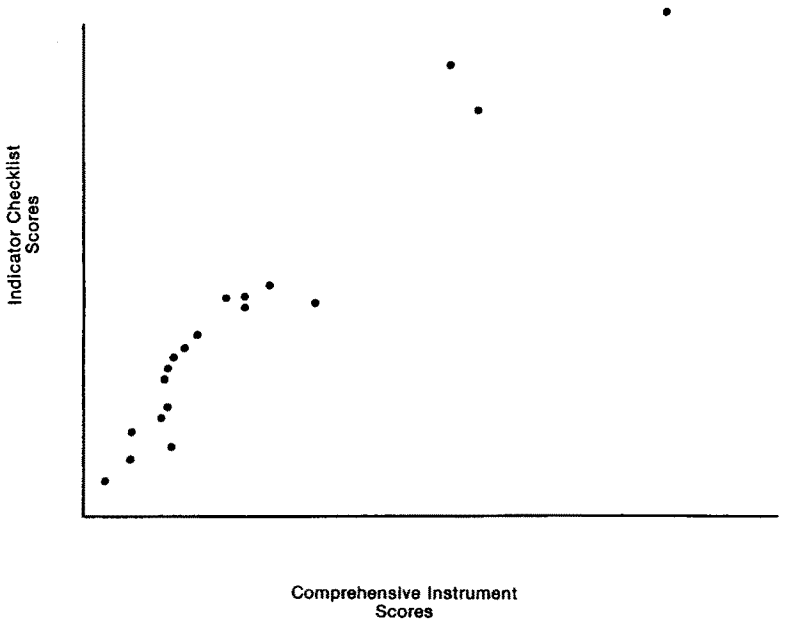


FIGURE 2a

Based on the results of Pennsylvania's, West Virginia's, California's and New York City's Indicator Checklists, certain common items were consistently showing up as predictor items that were separating those good providers from those problem providers. In other words, the following items were always in compliance for the good providers and were always out of compliance for the problem providers:

LICENSING SUBSCALE

A. GROUP SIZE AND ADULT CHILD RATIOS;

INFANTS	1 STAFF TO 5 CHILDREN 10 INFANTS IN A GROUP
TODDLERS	1 STAFF TO 4 CHILDREN 8 TODDLERS IN A GROUP
PRESCHOOLERS	1 STAFF TO 10 CHILDREN 20 PRESCHOOLERS IN A GROUP
SCHOOL AGE	1 STAFF TO 15 CHILDREN 30 SCHOOL AGE CHILDREN IN A GROUP

B. SUFFICIENT SPACE—MINIMUM OF 40 SQ FT PER CHILD;
C. EQUIPMENT IS EASILY ACCESSIBLE TO CHILDREN;
D. ALL VEHICLES ARE EQUIPPED WITH AGE-APPROPRIATE SAFETY CARRIERS;
E. CLEANING MATERIALS ARE INACCESSIBLE TO CHILDREN;
F. EMERGENCY CONTACT INFORMATION IS AVAILABLE FOR ALL CHILDREN;
G. ALL STAFF HAVE HAD PERIODIC HEALTH APPRAISALS;
H. ACTIVITIES PROMOTE:
DEVELOPMENT OF SKILLS
SELF-ESTEEM
POSITIVE SELF-IDENTITY
CHOICE OF ACTIVITIES.
(Fiene, 1984)

To most administrators and policymakers, the advantages of a shorter form will be readily apparent. The short form extends the general advantages of an IPM system in three key ways.

First, it substantially reduces the burden on providers, especially those providers that have a record of high compliance and are judged

suitable for use of the short form—it is proposed that these providers be visited once every three years using the comprehensive instrument. In the intervening years, the indicator checklist should be used.

Second, the indicator checklist approach can further reduce a state's cost of monitoring and permit the more efficient reallocation of staff resources to other activities. A cost effectiveness study conducted in West Virginia utilizing their indicator checklist resulted in a savings of 50% staff time in determining the level of compliance of providers (in dollars, this translated to $800 annually per visit saved (Peat, Marwick, & Mitchell 1983). With such a substantial savings in time, program monitors/evaluators could be freed to act more as consultants in providing technical assistance to providers.

Third, reviews of providers may be consolidated where appropriate. For example, state staff who perform fiscal/contract compliance audits of providers might be trained to administer the indicator checklist during their audit.

The total effect of maintaining a strong compliance monitoring capability that is less of a burden on providers and that achieves greater efficiency with lower cost is a higher quality monitoring system.

What is Needed to Develop an Indicator Checklist?

An indicator checklist is constructed as follows (See Figure 3):

1) Begin with an existing, comprehensive instrument that has a sufficiently large number of items so as to make greater efficiency desirable. The relative importance of each item as reflected in some kind of scoring or weighting system must have been established. Many criteria may be used for weighting the individual items. One criterion that is particularly useful for weighting purposes is the extent to which a particular item is related to health, safety, or developmental risks to children.
2) Your state should have used the comprehensive instrument long enough so that it is considered reliable for monitoring purposes; the instrument should have generated data that can be used to distinguish among providers in substantial compliance and weak or non-compliant providers.
3) With an existing, comprehensive instrument and some historical score information, it is possible to use a simple arithmetical formula (phi coefficient) to select those items from the long questionnaire that are most useful in distinguishing be-

tween good and inadequate programs. These distinguishing or "predictor" items form the basis of the indicator checklist (See Fiene & Nixon, 1983) for a detailed explanation of the formula for developing an indicator checklist).

4) The final step is to include on the short form particular questions or items from the comprehensive instrument that are of critical importance to the health and safety of children. Typically, these are items which, if violated, would be sufficient basis for denying or revoking a license for a day care program. Usually, such items are few in number. They are added to the short form with the predictor items to ensure that children will not be jeopardized by any statistical errors that might occur if only the "predictor" items were used.

From this description of the procedure for developing the shortened instrument, it is clear that the essential prerequisites for such a checklist are: 1. a long, comprehensive instrument in which state administrators have confidence; 2. items on the comprehensive instrument that are weighted to indicate their relative importance; 3. sufficient score data from use of the comprehensive instrument to differentiate among better and worse programs; and 4. state commitment to developing a short form instrument.

Specific Concerns of Administrators and Policymakers

It may be useful to address particular concerns of administrators and policymakers who may be interested in or even actively considering developing a shortened form of their state's monitoring or

Constructing The Indicator Checklist

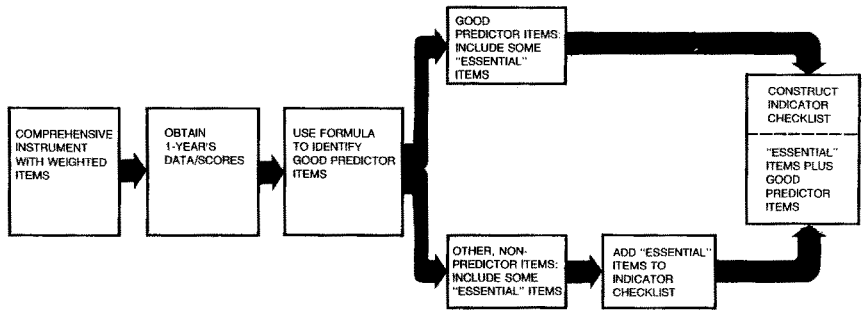


FIGURE 3

licensing questionnaire or checklist. In particular, administrators will need to know: how their state can make use of an indicator checklist; whether indicator checklists have been tried by other states; how the quality of monitoring can be ensured; and whether there are potential drawbacks.

Can My State Make Use Of An Indicator Checklist?

Practically every state that presently has some form of questionnaire or checklist can potentially profit from using a shortened form of the instrument. Naturally, if your state's instrument is already sufficiently short, then little will be gained by being more selective about questions or items to include. Many states are confronted, however, with lengthy instruments that cover a wide range of requirement areas. These states are prime candidates for short-form instruments.

Similarly, perhaps obviously, if your state does not currently have an instrument-based system, then consideration of an indicator checklist/short form is premature.

In order to develop a successful indicator checklist, it is important that the items on your state's current instrument be clearly linked to:

1. Your state's requirements (regulations); and
2. The results or outcomes that are considered desireable with respect to the providers' performance in such areas as licensing, contract monitoring, and program quality.

Unless there is a clear correspondence between intrument items and requirements, there is a danger that the items selected for inclusion on the short form will be only loosely tied to regulations and may be perceived by providers as improper or illegal. Similarly, if there is only a weak link between items on your state's comprehensive instrument and the results that you expect from providers, then the ground for selecting particular items as good predictors will not be solid enough.

Have Indicator Checklists Been Tried By Other States?

The concept of an indicator checklist may be appealing, but administrators are usually hesitant to take risks that could jeopardize systems that have been developed through years of work. It is often satisfying to know that other states have already tested the concept in practice.

At present, the indicator checklist concept is still an innovation that holds great promise but has been fully implemented in only four

states; Pennsylvania, West Virginia, New York, and California have developed an indicator checklist/short form and are testing the concept. Because the initial analyses conducted by these states suggest that the short form can work, other states such as Michigan and Texas have declared their intention to develop a shortened instrument by using these states' experiences as a guide. Clearly though, the indicator checklist/short-form methodology is still in the experimental stage.

How Can The Quality Of Monitoring Be Ensured?

Top administrators may wonder whether the shortened instrument presented here will compromise the quality of their state's current monitoring effort. Our view is that the short form will enhance current monitoring efforts by increasing the efficient and effective utilization of monitoring staff. But there are precautions that states should take in developing and using indicator checklists.

The indicator checklist/short instrument should not be used as a substitute for the comprehensive instrument, but rather as its complement. If the short form is viewed as the monitoring instrument, then there may be a tendency over time for providers to meet only the requirements covered on the short form. This situation could, indeed, compromise the quality of monitoring.

On the contrary, we would anticipate that states might keep their comprehensive instruments as the definitive set of compliance expectations and administer them for the initial review (e.g., licensing review) of a provider, and could use the indicator checklist/short form as:

1. A screening device to determine whether, for a given provider, it is necessary to administer the longer version; and
2. An interim review instrument to be used as the principal tool for providers who have a good record of compliance.

For example, the comprehensive instrument might continue to be used for "problem" providers and on a periodic basis, say, every three years for good providers. Naturally, if the short form were used with a provider and problems were discovered, then the comprehensive instrument, or some portions of it, could be administered.

Over time, as conditions change, it will be necessary to update and revise both the comprehensive and short instrument. Using the comprehensive instrument at least periodically with all providers will provide a basis for modifying the short form to reflect changing compliance patterns.

We expect that both versions of the instrument would be used by state staff who are trained and competent to assess compliance. These staff would certainly not limit themselves to using the short form if they determined, on site, that conditions warranted using the comprehensive instrument. The purpose of the indicator checklist/short form is to increase the options available to the state for monitoring in a flexible and cost-effective manner, not to put unreasonable constraints or "blinders" on monitoring staff.

What Are The Potential Drawbacks?

As with all innovations, the introduction of an indicator checklist as the basis for routine monitoring in a state may create some problems. Because so few states have introduced indicator checklists on a widespread basis, it is difficult to identify all of the concerns that may arise in practice. However, a few potential problems can be anticipated. (See Table 2).

First, some states' regulations require that all providers be reviewed every year in all regulatory areas. That is, the state insists that a comprehensive review, for example, using the comprehensive form of a state's monitoring instrument, take place for each provider. If this is the situation in your state, then the use of a shortened instrument may depend on changing the current regulatory provisions concerning the frequency and scope of reviews. A strong basis for making such a change is the cost effectiveness of the indicator checklist/short form, that is, its potential for reducing monitoring costs substantially without reducing the quality of the monitoring effort.

TABLE 2

Potential Drawbacks	Possible Solutions
• Regulatory Requirement for Annual Comprehensive Review	• Change Regulatory Requirements
• Staff Resistance	• Educate Staff
• State's Lack of Prerequisites	• Seek Assistance in Obtaining Prerequisites

Second, the state's staff who are responsible for monitoring may resist the introduction of the indicator checklist/short form. From their viewpoint, it may appear that the use of indicator checklists is a reduction in the importance of their professional roles and that the

state's cost savings may take the form of fewer jobs for day care monitors.

In our view, states may need to assure their staff that the indicator checklist/short form is not intended to reduce either the professional judgments involved or the scope of the monitoring function. As mentioned earlier, the comprehensive and short instruments must be used in a complementary way, not as substitutes, in order for the short form to have validity. If anything, the judgment of the monitors may be expanded as it becomes necessary to decide whether, in a particular case, the short instrument will be sufficient to measure compliance with state requirements, and/or program quality criteria. Monitors must be persuaded that the short form is an aid that is designed to reduce the monitors' workload for those providers with whom the short form is appropriate.

The reduction in workload may gradually change the relationship of monitors to providers from one of regulation to one of active support in improving the health and safety of the day care environment and encouraging child development. This change in the monitors' role could enable the state to make even better use of the current monitoring staff's knowledge and experience.

With respect to costs and staff reduction, there is little question that substantial decreases in workload could also result in reduced staffing levels. However, before considering cutbacks in staff, we would encourage states to consider reallocating staff time that is saved because of the short form to other monitoring activities such as technical assistance to providers involving program quality issues.

Third, a state may discover that it does not have the necessary prerequisites, described earlier, to develop and implement an indicator checklist. If your state lacks these prerequisites—in particular a comprehensive instrument, reports of scores, and a system of weighting items on the instrument—then it may be advantageous for you to examine other reports prepared by the Children's Services Monitoring Transfer Consortium that describe how these prerequisites can be met. You may be interested in obtaining the Consortium's series of Guide Books. The three volumes of this series describe in detail how to develop a comprehensive instrument from which an indicator checklist/short form can be derived.

Conclusion

The art of monitoring has evolved considerably in recent years as more highly trained staff have been given responsibility for monitoring, and as clearer procedures, such as instrument-based program monitoring, have been implemented. This evolution has con-

tributed positively to achieving the desirable outcomes of improved day care for children for which the state has developed regulations. At the same time, the evolution has, we hope, made it possible for providers to operate more effectively with the minimum necessary oversight by the state.

Instrument Based Program Monitoring Systems are now being developed in other children's services such as MH/MR services. Pennsylvania has developed its child welfare information system based on the instrument based program monitoring concept. This system meets two needs for Pennsylvania: it tracks children through its foster care system; and it complies with PL 96-272—the Adoption Assistance and Foster Care Act—a federal law. West Virginia is attempting to use the IPM methodology in monitoring its family day care home programs.

Also, a micro-computer, decision support system based on the Instrument Based Program Monitoring and Indicator Checklist methodology is being developed by the Children's Services Monitoring Transfer Consortium (CSMTC). The CSMTC is a group of states (Pennsylvania, West Virginia, California, New York, Michigan, and Texas) who have been disseminating exemplary monitoring techniques from state to state. Based on the combined efforts of these states, a generic indicator checklist that measures compliance with state regulations as well as program quality has been developed (Fiene, 1984). The CSMTC feels that this generic indicator checklist can be used by states who have not developed an instrument to assess providers, or as a model instrument to assist states in developing their own instruments.

The real potential of monitoring in achieving social goals, (such as protecting the health and safety of young children, ensuring quality child development programs, and tying these to child development outcomes), will be better realized through continuing research and development of improved monitoring procedures. It is in this context that the development of the indicator checklist represents a major advance in monitoring children's services.

References

Aronson, S., Fiene, R. & Douglas, E. *Pennsylvania Child Development Program Evaluation Instrument,* Bureau of Child Development, Harrisburg, Pennsylvania, 1977.

Belsky, J.& Steinberg, L. The effects of day care: A critical review, *Child Development,* 1978, *49,* 929-949.

Class, N. & Orton, R. Day care regulation: The limits of licensing, *Young Children,* Vol. #2, 1980, September, 12-17.

Collins, R. Child care and the states: The comparative licensing study, *Young Children,* Vol. #2, 1983, July, 3-11.

Cohen, M. *New York City Indicator Checklist,* NYC Agency for Child Development, New York City, 1983.

CWLA, Day Care Standards Revisions, *Child Welfare League of America,* New York City, 1983.

Douglas, E. & Fiene, R. Making the almost impossible—possible: Evaluation of a human service (Ecological Paradigm for child care monitoring), paper presented at *National Association for the Education of Young Children,* November 1979, Atlanta, Georgia.

Ferrar, H., Gleason, D. & Smith, B. *A State-of-the-Art Report on State Monitoring of Child Care Facilities,* Social Services Research Institute, Washington, D.C., 1980.

Fiene, R., The two year old: Characteristics and management of his play, *Dimensions,* 1974, *2,* 2, 46-58.

Fiene, R., Current trends in research, *Children,* 1975, Spring, 15-18.

Fiene, R. *Child Development Program Evaluation Scale,* Children's Services Monitoring Consortium, Washington, D.C., 1984.

Fiene, R., Douglas, E., & Kroh, K. *Child Development Program Evaluation,* Bureau of Child Development, Harrisburg, Pa., 1977.

Fiene, R., Douglas, E. & Nixon, M. *System Documentation for Pennsylvania's Instrument Based Program Monitoring Information System,* Children's Services Monitoring Services Consortium, Washington, D.C., 1983.

Fiene, R. & Nixon, M. *Instrument Based Program Monitoring System,* Children's Services Monitoring Consortium, Washington, D.C., 1981.

Fiene, R. & Nixon, M. *Indicator Checklist Methodology,* Children's Services Monitoring Consortium, Washington, D.C., 1983.

Harms, T. & Clifford, R. *Early Childhood Environment Rating Scale,* Teachers College, Columbia University, New York, 1980.

HEW, *FIDCR Report of Findings and Recommendations,* U.S. Department of Health Education and Welfare, Washington, D.C., 1978.

Lazar, I. *Lasting Effects After Preschool,* U.S. Department of Health, Education and Welfare, Washington, D.C., 1979.

Merrill, B. *West Virginia Indicator Checklist,* West Virginia Department of Social Services, Charleston, W.V., 1981.

Morgan, G. *Regulation of Early Childhood Programs,* Washington, D.C.: Day Care and Child Development Council of America, Inc., 1971.

Morgan, G. *Alternatives for Regulation of Family Day Care Homes for Children,* Washington, D.C.: Day Care and Child Development Council of America, Inc., 1974.

Morgan, G. Regulating early childhood programs in the eighties, unpublished manuscript, 1980.

NAEYC, Criteria for quality early childhood programs, *National Association for the Education of Young Children,* Washington, D.C., 1983.

Peat, Marwick, & Mitchell, Assessment of the inplementation of the indicator checklist for day care centers in West Virginia, *Children's Services Monitoring Consortium,* September, 1983.

Poole, J. *Child Development Program Quality Review,* California State Department of Education, Sacramento, CA., 1982.

Ruopp, R. *Children at the Center,* Cambridge, MA: Abt Books, 1979.

Watson, S., Fiene, R., & Woods, L. *Kids MicroComputer Software,* Children's Services Monitoring Consortium, Washington, D.C., 1984.

Zigler, E. & Gordon, E. *Day Care: Scientific and Social Policy Issues,* Boston, Massachusetts: Auburn House Publishing Company, 1982.

Regulatory Compliance and Quality Programs: Constraints and Opportunities for Integration

Brian Freer[a1], PhD and Richard Fiene[b], PhD ORCID/0000-0001-6095-5085

[a]*Center for Science, Technology, Medicine and Society; University of California, Berkeley. Corresponding Author: bjfreer@berkeley.edu*

[b]*Edna Bennett Pierce Prevention Research Center, Pennsylvania State University.*

Abstract

Management systems for regulatory compliance and quality programs are examined in this paper from the standpoint of their potential integration and in terms of the concept of a process. The paper identifies five common drags on management system optimization and outlines a scoring system that organizations may use to evaluate their management systems for potential adoption of an integrated process-based program.

Key Words: Management system; implementation; integration; ISO; regulatory compliance

1. Introduction

This paper argues that by examining the costs and benefits of implementing process-based integrated management systems, organizations may gain insight into the potential value of merging regulatory compliance with quality programs. In this paper, we define regulatory compliance as all government requirements, exclusive of accounting, facing an organization and the activities an organization takes to conform to these requirements. Quality programs are defined as customer, international and national standards, and other requirements where an organization is obligated to show conformance. According to ReVelle (2003), a process is "a series of sequentially oriented, repeatable events that have both a beginning and an end, and which result in either a product or a service." Research by Carvalho et al. (2015) makes a strong case for integrating multiple standards (e.g., quality, environmental, safety) into a single management system. Their research findings are outlined further below. They define an integrated management system as "a set of interrelated processes that share human resources, information, materials, infrastructure, (sic) financial resources" (Carvalho et al. 2015). However, their research focuses exclusively on international standards, setting aside questions concerning regulatory compliance. Fiene (2019, 2022) has recently made the case that regulatory compliance programs may be enhanced by incorporating measurement and continual improvement—hallmarks of quality initiatives—into regulatory compliance programs. Our focus here is on the challenges and potential efficiencies that organizations may experience through

[1] Corresponding Author: Brian Freer - brian@freeconsulting.com

implementing a management system designed to merge regulatory compliance and quality into an integrated process-based management system.

Scholars have identified a lack of research on the topic of the relationship between quality programs and regulatory compliance (Doyle 2007; Doyle et al. 2014). Furthermore, researchers have noted that relatively few studies of implementation in regulatory compliance have been published to guide research (Panitz et al. 2011). Recent actions by the US Food and Drug Administration (FDA) presents an opportunity to highlight the relationship between regulatory compliance and quality programs in organizations.

On February 23, 2022, the FDA issued a proposed rule to align 21 CFR 820 (known as the Quality System Regulation) with ISO 13485: 2016. ISO 13485 is an international standard for medical device quality. The FDA's proposed rule intends to achieve this alignment by "incorporating by reference" ISO 13485 into 21 CFR 820. FDA's proposed rule to align its Quality System Regulation with ISO 13485 provides a catalyst to examine the relationship between regulatory compliance and quality with specific reference to implementation because of the process-based orientation of ISO 13485 and other ISO quality standards. Organizations looking to take advantage of the alignment of 21 CFR 820 with ISO 13485, and companies planning to explore management system integration of other regulations and standards, may not have process-based systems in-place to manage conformance. It is with these points in mind that this paper puts forth a basic framework that organizations may consider when assessing the costs and benefits of adopting an integrated process-based system.

This alignment of ISO 13485 and 21 CFR 820 is atypical and not yet conventionally found elsewhere when considering regulations and standards, however, organizations may take the initiative to create a process-based management system that integrates other regulatory compliance requirements with quality program standards. For example, the food processing industry is regulated by FDA's 21 CFR 117 known as Current Good Manufacturing Practice, Hazard Analysis, and Risk-Based Preventive Controls for Human Food (CGMP). An international standard in food processing quality is Safe Quality Food (SQF). An organization in this sector may face the dual requirements of FDA's CGMP and SQF's quality requirements (SQF Code Edition 9) and choose to integrate their programs to conform to both in a single management system. CGMP is not aligned with SQF by the regulator or standards publisher, respectively. However, an organization may utilize cross-reference matrices to integrate them into their management system. A cross-reference matrix shows shared requirements between two standards or regulations. A cross-reference matrix also lists requirements that are not shared, yet still required from one of the two standards or regulations under consideration.

Considering the background sketched out above, this move by the FDA to align a regulation with a process-based standard is unique because regulations are not typically written in a process framework, let alone aligned with a standard. To the contrary, regulations tend to be written in a policy narrative format listing requirements and end-state outcomes organizations must achieve. Process-based standards, on the other hand, are organized around the idea of allowing an organization the latitude to set their own metrics while requiring companies to show their reasoning for these targets while engaged in continual improvement, among other systematically related

activities. At the other end of the spectrum, a regulation is conventionally a set of inflexible rules that a government agency imposes on an organization. International, national, and industry standards differ from regulations in that they are not promulgated by government agencies, rather they are a set of technical specifications developed by an international body (e.g., International Organization for Standardization - ISO), a national standards entity (e.g., American National Standards Institute - ANSI) or an industry standard (e.g., American Institute of Steel Construction - AISC). Standards are typically adopted by organizations voluntarily, however organizations sometimes encounter customer requirements that stipulate conformance with a standard.

2. Common Problems Encountered by Organizations in Management System Implementation

Organizations inevitably encounter through external audits, executive reviews, and operations, inter-relationships between regulatory compliance and quality programs. However, research suggests that organizations do not generally prioritize investment in the design and implementation of management systems focused on controlling and optimizing this inter-relationship (Doyle 2007). Thus, outside of the realm of operations the relationship between compliance and quality in organizations is more often reacted to in a haphazard manner as opposed to intentionally integrating the two in a management system. As Doyle (2007) has pointed out, one reason for this predicament is that it is inherently difficult for organizations to coordinate legal and supra-legal requirements. Supra-legal refers to binding requirements faced by an organization in addition to government rules. Supra-legal requirements may include national association standards and customer specifications. Additionally, research has identified other roadblocks that deter organizations from integrating quality programs with regulatory compliance, including limited resources, lack of top management support, and inherent complexity (Doyle et al 2014). Scholars, with the notable exceptions cited above, generally treat the two topics separately. Government agencies and standards registrars, for their part, have historically avoided prescribing structural/organizational requirements concerning documents and their information format in management systems. In part due to this, government and certifying body auditors encounter a myriad of information management schemes in stand-alone and integrated management systems.

As seasoned management consultants can attest through their experiences encountering legacy management systems in organizations, it is common to find a set of characteristics that constrain organizational effectiveness in the pursuits of quality programs and regulatory compliance. Here we identify five constraints, based on decades of practice in the implementation of management systems for organizations through consulting. Each of these constraints are drags on the optimization of management system implementation.

First, organizations often react passively to externally generated regulatory compliance targets and accept them at face value as published by government agencies. When organizations accept targets at face value, the wider context and purpose of collecting and reporting data on a given topic may be ignored by an organization. Accepting compliance targets as-is may decrease the chances an organization has to undertake initiatives to explore data collection and reporting that are of value to the

organization, beyond just satisfying regulatory compliance.

Second, as a reaction to agency generated compliance targets, organizations may develop policies, procedures, forms and reports that are binary yes/no in format (e.g., was the target met?). Implementing a management system based on binary values stymies measurement.

Third, facing an array of compliance targets imposed externally, organizations may then decide to maintain two separate management systems, one for regulatory compliance and the other for their quality program. As a result of this, management systems can fall prey to becoming centralized silos of information. For the regulatory compliance system, but also encountered in quality programs, organizations may adopt an information management approach based on the sequence, numbering arrangement, and official language of the regulation and/or standard (i.e., an elements-based system). Adopting the language of an outside entity wholesale increases the chances that an organization will silo regulatory compliance information, thereby disconnecting this knowledge from the wider organization. Government agencies promulgate regulations, does it make sense for an organization with a unique culture and practices to follow a structure imposed from outside?

Fourth, organizations may create a narrative structure (text rather than process flowcharts and process maps) to carry out an elements-based system. Best practices in industry have moved away from narrative-based procedures in management systems because dense text is hard to follow; text-based policies are less likely to be linked to other activities in regular workflows than other graphical devices.

Best practices now utilize process flowcharts and process maps. Products such as Visio, along with an evolving world of web-based flowcharting tools, exist as resources. ISO 5807, the standard for flowchart symbols and methodology, is a helpful reference. ISO 5807 (1985: 1-2) identifies five types of flowcharts: data flowchart, program flowchart, system flowchart, program network chart, and system resource chart (see Figures 1 and 2 below).

ReVelle's (2003) definition of process flowchart is useful for the purposes of management system implementation explored in this paper. According to ReVelle (2003) a process flowchart is a graphical representation of a single process, using symbols to show the sequence of steps, typically moving left to right. Although the idea of a process map is not referenced in ISO 5807 because it came into use in industry after the standard was originally published in 1985, the marketplace adopted the concept because it nicely illustrates how processes are linked to one another. As ReVelle (2003) notes, a process map is "a two-dimensional version of a process flowchart that also portrays handoffs and receipts of products and/or services from one person, organization and/or location to another. A process map shows process inputs and outputs moving left to right but then connects to other processes sequentially by linking to subsequent processes in a top to bottom arrangement.

Fifth, by default narrative-based management systems are commonly structured on departmental organizational charts, rather than being based on individual process-ownership. Responsibility in the departmental organizational chart method rests with the department. This method means it is not clear who is responsible for

implementation and execution of a specific area. Additionally, with an organizationally based approach it is expensive and time consuming to change the structure of a management system every time an org chart changes. Allocating responsibility and authority at the individual level through process-ownership avoids the wasted effort of re-designing the management system when an organization changes structure. Furthermore, vesting process-ownership at an individual level makes it easy to locate the person responsible for a given process to obtain information and discuss improvements.

3. Characterizing Management System Attributes in an Organization

An organization may wish to evaluate its existing management system or explore options concerning implementing a new management system. This section, also based on practice in management consulting, applies to an organization seeking to better understand its current management system. This section, also based on practice in management consulting, applies to an organization seeking to better understand its current management system. A management system is defined as an information management framework that describes how an organization conforms to legal and supra-legal requirements concerning quality, environmental, and other aspects.

Organizations can undertake two activities that will provide a basis to outline the pros and cons of changing the structure of a management system or improving components of a management system. The first step is to identify all legal and supra-legal requirements facing an organization. Second, an organization's management system is categorized into the following types: elements-based or organized using an independent system; composed primarily with narrative, text-based procedures or process oriented with process flowcharts and process maps.

To identify legal requirements facing an organization, a table is generated containing rows of all known regulations that apply to the company within the scope the organization wishes to control. Next, all compliance points are detailed in a column adjacent to each regulation, including any required training along with written plans and/or procedures. Finally, the required records and reports are identified in an additional column. The same steps are undertaken by the company to identify supra-legal requirements. In the case of supra-legal requirements, an organization should identify international standards, national standards, contractual customer requirements, corporate policies, insurance requirements, and trade association standards.

With the legal and supra-legal requirements in a table, the organization may proceed to characterize the format of its existing management system. This step begins with an understanding of the typical components of a management system. Management systems often contain a brief manual at the front that spells out the scope of the management system by listing the regulations and standards that the system covers along with a related scope of operations. Procedures and work instructions, the how-to of the management system, typically follow the manual. Finally, forms and records round out the management system. To assess the current state of the management system an organization should know the overall structure of the information and the type of format it is using for procedures and work instructions. There are two common types of general information management structures organizations use in management systems:

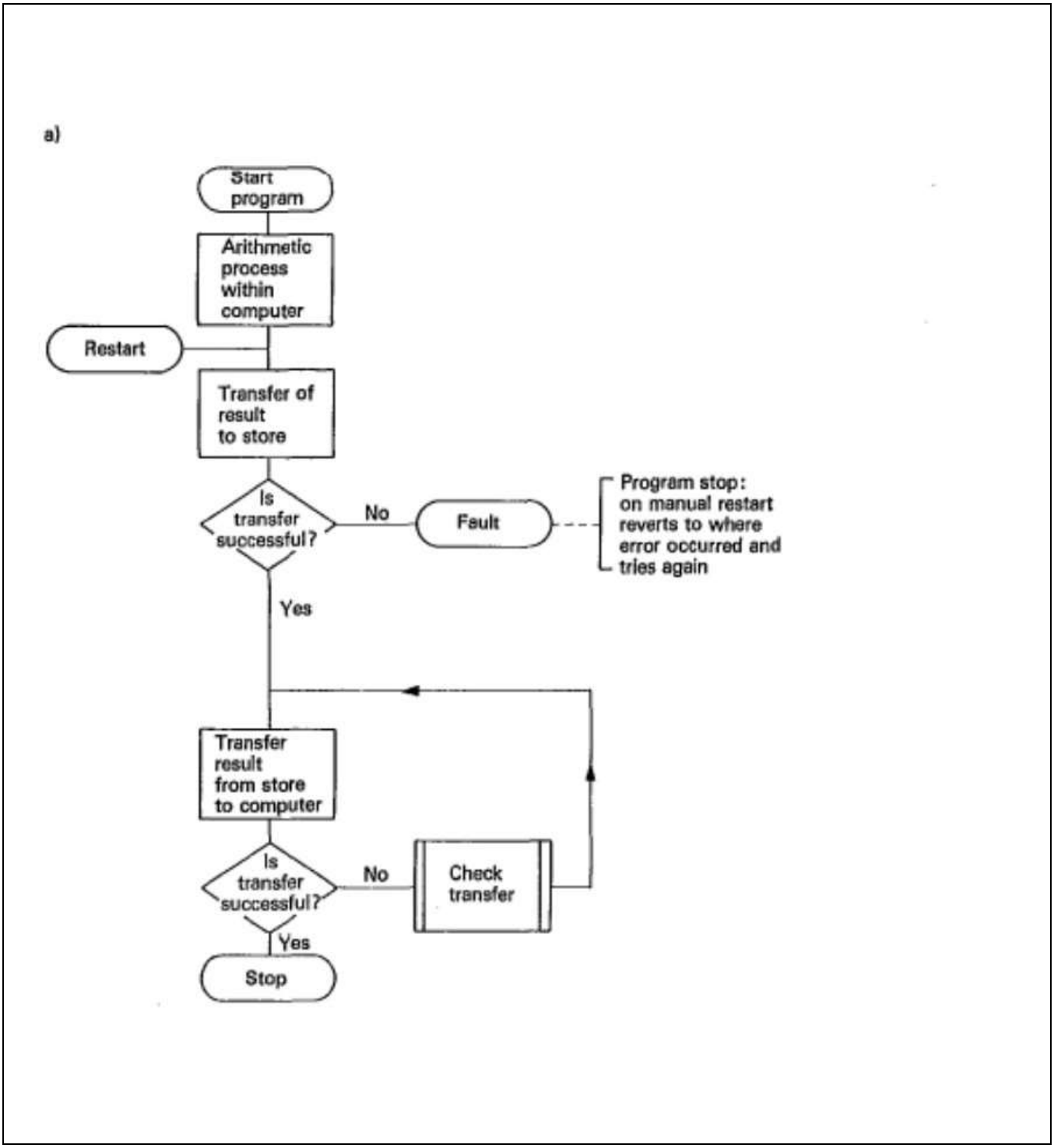


Figure 1: Program Flowchart, from ISO 5087 Annex B

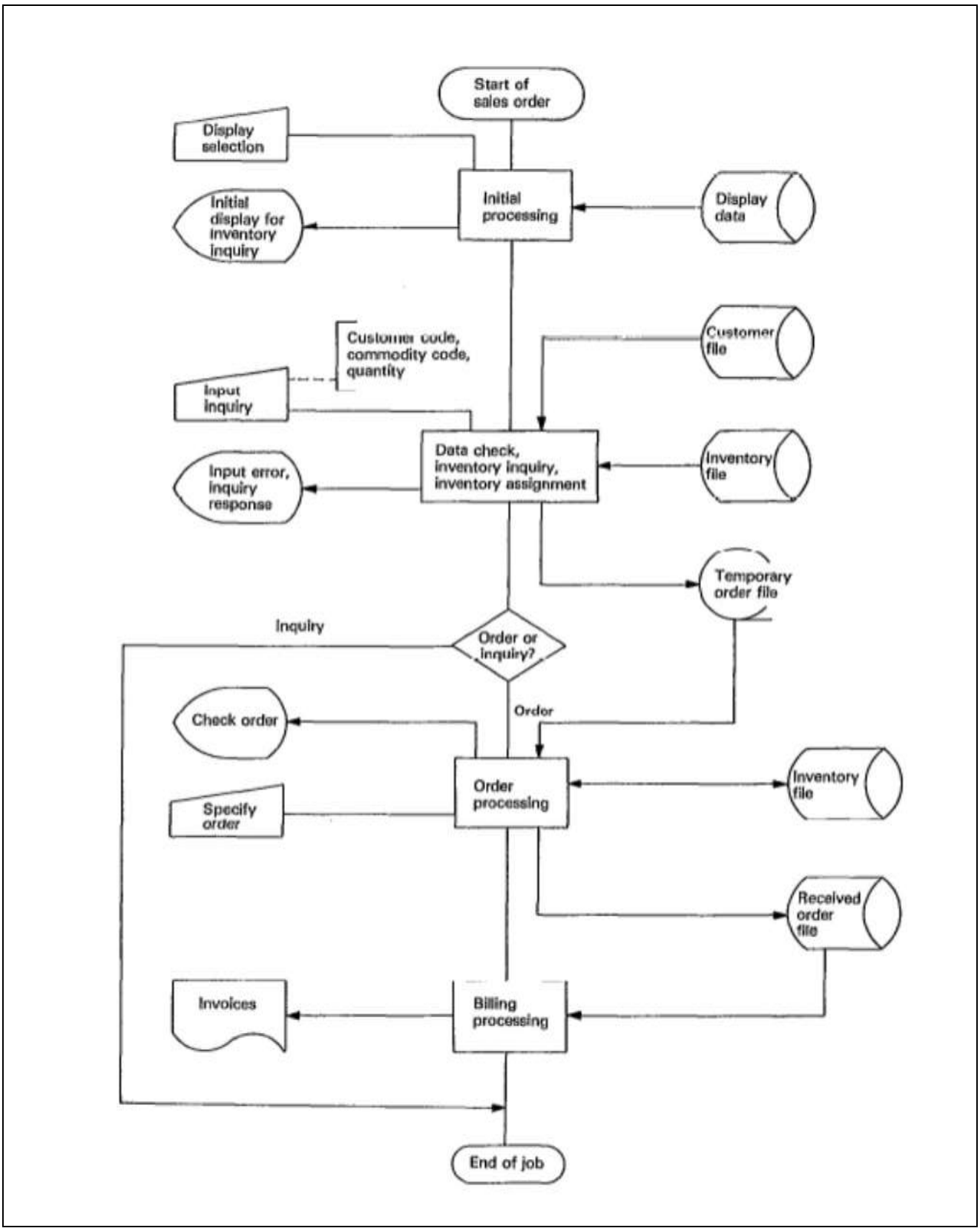


Figure 2: System Flowchart, from ISO 5087, Annex C

elements-based and independently structured. An elements-based system follows the numbering and order of the regulation and/or standard. An independently structured system is based on a generic ordering using language common to the organization.

To characterize the format of procedures, begin by selecting three procedures from separate areas of the management system. Read the procedure and look for one of two common possibilities: narrative sentences organized in statements or graphical depictions of the steps of an activity (process flowcharts and process maps). If no process flowcharts or process maps are encountered, the organization has a narrative management system. If the organization's management system contains process flowcharts and process maps, the possibility exists that the management system is process-based.

With the table of legal and supra-legal requirements in hand along with the findings of the assessment of the management system format, an organization can next score their findings. To score the findings, begin by reviewing the table of requirements. If the table contains many requirements and the majority of these are complicated, then issue a score of High/Complex. The table's listings may be scored Medium/Standard if the organization is not in a highly regulated sector. Finally, a score of Low/Simple may be assigned to companies that are lightly regulated.

To score the overall structure of the management system, assign a label of Elements for systems that follow the sequence and nomenclature of the regulation and/or standard. If the system is based on the organization's own approach, label it independent. For procedures, assign a value of Primarily Narrative for management systems where most of the information in the procedures and work instructions is in sentences of text without graphical flowcharts. If an organization encounters a management system where the procedures and work instruction are mainly composed of flowcharts and maps, assign a value of Primarily Process.

Organizations with a combined score of High/Complex, Elements, and Primarily Narrative may find value in considering a transition to a process-based management system.

4. Challenges and Advantages to Implementing a Process Oriented Management System Integrating Compliance and Quality in an Organization

This paper concludes by briefly identifying some challenges and advantages organizations may encounter in the transition to a process-based integrated management system.

There are two primary challenges to an organization seeking to transition to a process-based approach to its management system. First, looking at each area of an organization as a series of activities characterized by an input and an output may be new to employees. A process-based approach also requires a shift in mindset for team members with no previous experience with a process-based system. Second, an organization moving to a process-based management system should plan for an activity that often requires six months to a year to accomplish. The transition to a process-based management system, the implementation phase, can be time consuming.

There are three main advantages of adopting a process-based management system. First, this approach facilitates the establishment of a baseline. Second, with a process baseline established, a company may then expend

less effort to set metrics and measure against a baseline. Third, the activity of continual improvement is enhanced through a process approach because an organization has established benchmarks for each process.

In their study of management system integration, Carvalho et al. (2015) found that the primary barrier to integration is a lack of collaboration between managers in the different areas (e.g., quality, environmental, safety). In the case of the integration of quality programs and regulatory compliance the nature of the relationship between managers in these areas would likely be a key factor in project implementation.

Among the benefits of integration outlined by Carvalho et al. (2015), six findings may be useful for organizations to consider. First, an integrated system uses a shared resources approach so there is only one procedure for auditing, purchasing, and corrective-action, for example. Second, team members found that it was easier to manage a single system. Third, it took less time to audit the system. Fourth, the organization experienced less time spent in meetings. Fifth, there was increased understanding of the entire system. Sixth, the organization may experience reduced costs.

In summary, the case for process-based management system integration is organizational efficiency, elimination of redundancy, along with compliance improvement through enhanced knowledge and measurement brought about by feedback from quality initiatives.

5. References

[1] Carvalho, K., Picci, F., Camarini, G., & Chamon (2015). Benefits in the implementation of safety, health, environmental and quality integrated system. *International Journal of Engineering and Technology* 7, no. 4, 355.

[2] Doyle, E. (2007). Compliance obstacles to competitiveness. *Corporate Governance: The international journal of business in society*, 612-622.

[3] Doyle, E., McGovern, D., & McCarthy, S. (2014). Compliance–innovation: integrating quality and compliance knowledge and practice. *Total Quality Management & Business Excellence* 25, no. 9-10, 1156-1170.

[4] Fiene, R. (2022). Regulatory Compliance Monitoring Paradigms and the Relationship of Regulatory Compliance/Licensing with Program Quality: A Policy Commentary. *Journal of Regulatory Science* 10, no. 1, 1-7.

[5] Fiene, R. (2019). A treatise on theory of regulatory compliance, *Journal of Regulatory Science*, 7, no. 1, 1-3.

[6] International Organization for Standardization. (1985). Information processing – documentation symbols and conventions for data, program and system flowcharts, program network charts and system resources charts. (ISO Standard No. 5807: 1985).

[7] Panitz, J. C., Wiener, M., & Amberg, M. (2011) Factors facilitating compliance implementation – Case study results from multinational enterprises, ECIS 2011 Proceedings, 3.

[8] ReVelle, J. B. (2003). Safety process analysis. *Professional Safety* 48, no. 7, 19-26.

Regulatory Compliance Monitoring Paradigms and the Relationship of Regulatory Compliance/Licensing with Program Quality: A Policy Commentary

Richard Fiene, PhD [a,*]

[a] *Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University*

Abstract

This policy commentary deals with two key issues within regulatory science related to the best methods for measuring regulatory compliance: Program monitoring paradigms and the relationship of regulatory compliance/licensing with program quality. Examples from program monitoring paradigms include: 1) Substantial versus Monolithic. 2) Differential Monitoring versus One size fits all monitoring. 3) "Not all standards are created equal" versus "All standards are created equal". 4) "Do things well" versus "Do no harm". 5) Strength based versus Deficit based. 6) Formative versus Summative. 7) Program Quality versus Program Compliance. 8) 100-0 scoring versus 100 or 0 scoring. 9) QRIS versus Licensing. 10) Non-Linear versus Linear. Examples from the relationship of regulatory compliance/licensing with program quality include: 1) "Do no harm" versus "Do good". 2) Closed system versus Open system. 3) Rules versus Indicators. 4) Nominal versus Ordinal measurement. 5) Full versus Partial compliance. 6) Ceiling effect versus No Ceiling effect. 7) Gatekeeper versus Enabler. 8) Risk versus Performance.

Keywords: regulatory compliance, program monitoring, licensing, program quality.

Introduction

This commentary on policy will deal with two key issues within regulatory science that need to be dealt with by licensing researchers and regulatory scientists as they think through the best methods for measuring regulatory compliance: 1) Program monitoring paradigms; 2) Relationship of regulatory compliance/licensing and program quality. The examples drawn are from early childcare and education but the key elements and implications can be applied to any field of study related to regulatory science that involves rules/regulations/standards. For the purposes of this manuscript "rules" will be used to

* Corresponding author: Richard Fiene; Email: rjf8@psu.edu;
Phone: 717-598-8908; ORCID: http://ORCID: 0000-0001- 6095-5085.

describe or refer to "rules/regulations/standards".

Program Monitoring Paradigms:

This section provides some key elements to two potential regulatory compliance monitoring paradigms (Differential/Relative versus Absolute/Full) for regulatory science based upon the Regulatory Compliance Theory of Diminishing Returns (Fiene, 2019).

As one will see, there is a need within regulatory science to get at the key measurement issues and essence of what is meant by regulatory compliance. There are some general principles that need to be dealt with such as the differences between individual rules and rules in the aggregate. Rules in the aggregate are not equal to the sum of all rules because all rules are not created nor administered equally. And all rules are to be adhered to, but there are certain rules that are more important than others and need to be adhered to all the time. Less important rules can be in substantial compliance most of the time but important rules must be in full compliance all of the time (Fiene, 2019).

Rules are everywhere. They are part of the human services landscape, economics, banking, sports, religion, transportation, housing, etc... Wherever one looks we are governed by rules in one form or another. ***The key is determining an effective and efficient modality for negotiating the path of least resistance in complying with a given set of rules***[2]. It is never about more or less rules, it is about which rules are really productive and which are not. Too many rules stifle creativity, but too few rules lead to chaos. ***Determining the balance of rules is the goal and solution of any regulatory science paradigm.***

Differential/Relative versus Absolute/Full Regulatory Compliance Paradigms: this is an important key organizational element in how rules are viewed when it comes to compliance. For example, in an absolute/full approach to regulatory compliance either a rule is in full compliance or not in full compliance. There is no middle ground. It is black or white, no shades of gray as are the cases in a differential/relative paradigm. It is 100% or zero. In defining and viewing these two paradigms, this dichotomy is the organizational key element for this paper. In a differential/relative regulatory compliance paradigm full compliance is not required and emphasis on substantial regulatory compliance becomes the norm.

Based upon this distinction between differential/relative and absolute/full regulatory compliance paradigms, what are some of the implications in utilizing these two respective approaches. Listed below are the basic implications that occur when selecting either of the two approaches on program monitoring systems: differential/relative versus absolute/full regulatory compliance paradigms.

There are ten basic implications that will be addressed: 1) Substantial versus Monolithic. 2) Differential Monitoring versus One size fits all monitoring. 3) "Not all standards are created equal" versus "All standards are created equal". 4) "Do things well" versus "Do no harm". 5) Strength based versus Deficit based. 6) Formative versus Summative. 7) Program Quality versus Program Compliance. 8) 100-0 scoring versus 100 or 0 scoring. 9) QRIS versus Licensing. 10) Non-Linear versus Linear.

1) Substantial versus Monolithic: in monolithic regulatory compliance monitoring systems, it is one size fits all, everyone gets the same type of review (this is addressed in the next key element below) and is more typical of an absolute paradigm orientation. In a substantial regulatory compliance monitoring system, programs are monitored on the basis of their past compliance history and this is more typical of a relative paradigm orientation. Those with high compliance may have fewer and more abbreviated visits/reviews while those with low compliance have more comprehensive visits/reviews.

2) Differential Monitoring versus One Size Fits All Monitoring: how does this actually look in a program monitoring system. In differential monitoring (Differential/Relative Paradigm), more targeted or focused visits are utilized spending more time and resources with those problem programs and less time and resources with those programs that are exceptional. In the One Size Fits All Monitoring (Absolute/Full Paradigm), all programs get the same type/level of review/visit regardless of past performance.

3) “Not all standards are created equal” versus “All standards are created equal”: when looking at standards/rules/regulations it is clear that certain ones have more of an impact on outcomes than others. For example, not having a form signed versus having proper supervision of clients demonstrates this difference. It could be argued that supervision is much more important to the health and safety of clients than if a form isn’t signed by a loved one. In a differential/relative paradigm, all standards are not created nor administered equally; while in an absolute/full paradigm of regulatory compliance, the standards are considered created equally and administered equally.

4) “Do things well” versus “Do no harm” (this element is dealt with in the second component to this paper below as well): “doing things well” (Differential/Relative Paradigm) focuses on quality of services rather than “doing no harm” (Absolute/Full Paradigm) which focuses on protecting health and safety. Both are important in any regulatory compliance monitoring system but a balance between the two needs to be found. Erring on one side of the equation or the other is not in the best interest of client outcomes. "Doing no harm" focus is on the "least common denominator" – the design and implementation of a monitoring system from the perspective of focusing on only 5% of the non-optimal programs ("doing no harm") rather than the 95% of the programs that are "doing things well".

5) Strength based versus Deficit based: in a strength-based monitoring system, one looks at the glass as “half full” rather than as “half empty” (deficit-based monitoring system). Emphasis is on what the programs are doing correctly rather than their non-compliance with standards. A strength-based system is non-punitive and is not interested in catching programs not doing well. It is about exemplars, about excellent models where everyone is brought up to a new higher level of quality care.

6) Formative versus Summative: differential/relative regulatory compliance monitoring systems are formative in nature where there is an emphasis on constant quality improvement and getting better. In absolute/full regulatory compliance monitoring systems, the emphasis is on being the gate-keeper (more about the gate-keeper function in

the next section on regulatory compliance/licensing and program quality) and making sure that decisions can be made to either grant or deny a license to operate. It is about keeping non-optimal programs from operating.

7) Program Quality versus Program Compliance: (this element is dealt with in greater detail in the second component of this manuscript) differential/relative regulatory compliance monitoring systems focus is on program quality and quality improvement while in absolute/full regulatory compliance monitoring systems the focus in on program compliance with rules/regulations with the emphasis on full, 100% compliance.

8) "100 – 0 scoring" versus "100 or 0 scoring": in a differential/relative regulatory compliance monitoring system, a 100 through zero (0) scoring can be used where there are gradients in the scoring, such as partial compliance scores. In an absolute/full regulatory compliance monitoring system, a 100% or zero (0) scoring is used demonstrating that either the standard/rule/regulation is fully complied with or not complied with at all (the differences between nominal and ordinal measurement is dealt with in the next section on regulatory compliance/licensing and program quality).

9) QRIS versus Licensing: examples of a differential/relative regulatory compliance monitoring system would be QRIS – Quality Rating and Improvement Systems. Absolute/full regulatory compliance systems would be state licensing systems. Many programs talk about the punitive aspects of the present human services licensing and monitoring system and its lack of focus on the program quality aspects in local programs. One should not be surprised by this because in any regulatory compliance system the focus is on "doing no harm" rather than "doing things well". It has been and continues to be the focus of licensing and regulations in the USA. The reason QRIS - Quality Rating and Improvement Systems developed in early care and education was to focus more on "doing things well" rather than "doing no harm". This is not the case in many Canadian Provinces and European countries in which they have incorporated program quality along with specific regulatory requirements.

10) Non-Linear versus Linear: the assumption in both differential/relative and absolute/full regulatory compliance monitoring systems is that the data are linear in nature which means that as compliance with rules increases positive outcomes for clients increases as well. The problem is the empirical data does not support this conclusion. It appears from the data that the relationship is more non-linear where there is a plateau effect with regulatory compliance in which client outcomes increase until substantial compliance is reached but doesn't continue to increase beyond this level. There appears to be a "sweet spot" or balancing of key rules that predict client outcomes more effectively than 100% or full compliance with all rules – this is the essence of the Theory of Regulatory Compliance (Fiene, 2019) – substantial compliance with all standards or full compliance with a select group of standards that predict overall substantial compliance and/or positive client outcomes.

As the regulatory science and administrative fields in general continue to think about the appropriate monitoring systems to be designed and implemented, the above structure should

help in thinking through what these measurement systems' key elements should be. Both paradigms are important, contexts, but a proper balance between the two is probably the best approach in designing regulatory compliance monitoring systems.

Regulatory Compliance/Licensing and Quality

This part of the policy commentary will delineate the differences between regulatory compliance and quality. It will provide the essential principles and elements that clearly demonstrate the differences and their potential impact on program monitoring. Obviously, there is some overlap between this section and the above section dealing with regulatory compliance monitoring paradigms. When we think about regulatory compliance measurement, we are discussing licensing systems. When we think about quality, we are discussing Quality Rating and Improvement Systems (QRIS), accreditation, professional development, or one of the myriad quality assessment tools, such as the Classroom Assessment Scoring System (CLASS) or Environment Rating Scales (ERS's). All these systems have been designed to help improve the health and safety of programs (licensing) to building more environmental quality (ERS), positive interactions amongst teachers and children (CLASS), enhancing quality standards (QRIS, accreditation), or enhancing teacher skills (professional development).

There are eight basic principles or elements to be presented (they are presented in a binary fashion demonstrating differences): 1) "Do no harm" versus "Do good". 2) Closed system versus Open system. 3) Rules versus Indicators. 4) Nominal versus Ordinal measurement. 5) Full versus Partial compliance. 6) Ceiling effect versus No Ceiling effect. 7) Gatekeeper versus Enabler. 8) Risk versus Performance.

1) Let's start with the first principal element building off what was discussed in the above section, "Do No Harm" versus "Do Good". In licensing, the philosophy is to do no harm, its emphasis is on prevention, to reduce risk to children in a particular setting. There is a good deal of emphasis on health and safety and not so much on developmentally appropriate programming. In the quality systems, such as QRIS, accreditation, professional development, Environmental Rating Scales, CLASS, the philosophy is to do good, its emphasis is looking at all the positive aspects of a setting. There is a good deal of emphasis on improving the programming that the children are exposed to or increasing the skill set of teachers or improving the overall environment or interaction that children are exposed to.

2) Closed system versus Open system. Licensing is basically a closed system. It has an upper limit with full compliance (100%) with all rules. The goal is to have all programs fully comply with all rules. However, the value of this assumption has been challenged over the years with the introduction of the Regulatory Compliance Theory of Diminishing Returns (Fiene, 2019). With quality systems, they tend to be more open and far reaching where attaining a perfect score is very difficult to come by. The majority of programs are more normally distributed where with licensing rules

the majority of programs are skewed positively in either substantial or full compliance. It is far more difficult to distinguish between the best programs and the mediocre programs within licensing but more successful in quality systems.

3) Rules versus Indicators/Best Practices. Licensing systems are based around specific standards/rules/regulations that either are in compliance or out of compliance. It is either a program is in compliance or out of compliance with the specific rule. With quality systems, there is more emphasis on indicators or best practices that are measured a bit more broadly and deal more with process than structure which is the case with licensing. It is the difference between hard and soft data as many legal counsels term it. There is greater flexibility in quality systems. With this said, if we can look at other service types, such as adult-residential services, there has been some limited success with blending structural and process elements but it still remains a measurement issue on the process side.

4) Nominal versus Ordinal measurement[3]. Licensing systems are nominally based measurement systems. Either you are in compliance or out of compliance. Nothing in-between. It is either a yes or no response for each rule. No maybe or partial compliance. With quality systems, they are generally measured on an ordinal level or a Likert scale. They may run from 1 to 3, or 1 to 5, or 1 to 7. There are more chances for variability in the data than in licensing which has 1 or 0 response. This increases the robustness of the data distribution with ordinal measurement.

5) Full or None versus Gradients or Gray Area. Building off of the fourth element, licensing scoring is either full or not. As suggested in the above elements, there is no in-between category, no gradient or gray area. This is definitely not the case with quality systems in which there are gradients and substantial gray areas. Each best practice can be measured on a Likert scale with subtle gradients in improving the overall practice.

6) Ceiling effect versus No Ceiling. With licensing there is definitely a ceiling effect because of the emphasis on full 100% compliance with all rules. That is the goal of a licensing program, to have full compliance. With quality systems, it is more open ended in which a ceiling effect is not present. Programs have many ways to attain excellence.

7) Gatekeeper versus Enabler: Licensing has always been called a gatekeeper system. It is the entry way to providing care, to providing services. It is a mandatory system in which all programs need to be licensed to operate. In Quality systems, these are voluntary systems. A program chooses to participate, there is no mandate to participate. It is more enabling for programs building upon successes. There are enhancements in many cases.

8) Risk versus Performance: Licensing systems are based upon mitigating or reducing risks to children when in out of home care. Quality systems are based upon performance and excellence where this is rewarded in their particular scoring by the addition of a new Star level or a Digital Badge or an Accreditation Certificate.

There has been a great deal of discussion in the early care and education field about the relationship between licensing, accreditation, QRIS, professional development, and technical

assistance. It is important as we continue this discussion to pay attention to the key elements and principles in how licensing and these quality systems are the same and different in their emphases and goals, and about the implications of particular program monitoring paradigms and measurement strategies. For other regulatory systems outside the human services field, the same type of model can be applied positioning compliance and quality as a continuum one building off of the other because I feel that with the introduction of more quality into a regulatory context will help to ameliorate the ceiling and plateau effect of diminishing returns on performance and outcomes.

Reference:

Fiene, R. (2019). A Treatise on Regulatory Compliance. *Journal of Regulatory Science*, Volume 7, 2019

Notes:

1. This manuscript should be read along with *A Treatise on Regulatory Compliance* which is referenced above because the two articles build off one another. In the *treatise* description, the specific idiosyncrasies of regulatory compliance data and other key implications of the theory are pointed out that enhance the presentation in this article, such as the extreme nature of skewness that is present in regulatory compliance data, nominal data measurement, the differences between full and substantial regulatory compliance, designing the most cost effective and efficient differential monitoring system, and the need to dichotomize data because of the skewed nature of the data distribution.

2. The ultimate goal is the most cost effective and efficient differential monitoring system for negotiating the path of least resistance in complying with a given set of rules which will provide the proper balance of rules. This should be the goal of any regulatory science paradigm. By using the previous *Treatise* article along with this article should provide a blueprint for the regulatory science field in designing a program monitoring system to measure regulatory compliance where an emphasis on differential monitoring should occur in licensing systems and full-scale monitoring should occur in program quality systems. Another approach is to have both regulatory compliance and program quality built as a continuum in the program monitoring system similar to what Head Start is attempting.

3. There are instances in which this dichotomy is not as clear or straightforward where licensing systems do allow partial compliance as a facility has opportunities to correct non-compliances on their way to achieving full compliance with specific rules. The problem is that this is not necessarily a standardized process and it is difficult to determine if it is used often in licensing agencies' monitoring efforts.

TDL.ORG TEXAS DIGITAL LIBRARY

Journal of Regulatory Science

http:\\journalofregulatoryscience.org

Regulatory Science

Journal of Regulatory Science 7 (2019) 1–3

A Treatise on the Theory of Regulatory Compliance

Richard J. Fiene[a,*]

[a]*Edna Bennett Pierce Prevention Research Center, Pennsylvania State University, 305 Templar Drive, Elizabethtown, Pennsylvania 17022*

Abstract

This treatise provides some insights into certain assumptions related to regulatory compliance and the implications for regulatory researchers and policy-makers for the future development of rules and regulations. Once regulatory compliance decision making moves from requiring full compliance with all rules to a substantial regulatory compliance decision making approach, the measurement and monitoring systems employed to assess programs and facilities change dramatically.

Keywords: regulatory compliance, risk assessment, key indicators, licensing, monitoring, measurement

1. Introduction

Regulatory compliance is a sub-discipline within regulatory science that focuses on measurement, monitoring systems, risk assessment, and decision making based on regulatory compliance scoring. Regulatory compliance is dominated by nominal scale measurement, that is, either a facility is in or out of compliance with specific rules. There is no middle ground with regulatory compliance as there is with most quality measurements, which are generally made on an ordinal scale. However, some regulators feel that certain regulations are not or should not be subjected to nominal measurement.

A factor with regulatory compliance data is that they generally follow a very skewed frequency distribution, which limits analyses to non-parametric statistics. Because of the skewed data distribution, dichotomization of data is warranted, given the lack of variance in the regulatory compliance frequency distribution - the majority of facilities [1] are either in full or substantial regulatory compliance.

An assumption within regulatory compliance is that full regulatory compliance, that is, 100 percent compliance with all rules [2], is the best (i.e., risk is minimized) possible scenario for the services being delivered and assessed. It is also assumed that all promulgated rules have an equal weight in their relative impact on the desired service delivery model, although this thinking has been changing over time regarding how rules are reviewed and complied with. This short treatise will examine the past 40 years of research delving into regulatory compliance measurement, and will provide some guidance to regulatory researchers and policy-makers as they move forward with both research and policy development related to rules. The data from these research studies have led to a Theory of Regulatory Compliance that demonstrates that substantial regulatory compliance - and not full regulatory compliance - is a more effective and efficient public policy as it relates to decision making on monitoring and licensing.

The results reported herein are drawn from human services delivery systems in the United States and Canada, such as early care and education, as well as child and adult residential services. The results are from state and provincial level licensing systems involving over 10,000 facilities serving over 100,000 clients. All the data are part of an international regulatory compliance database (https://data.mendeley.com/datasets/kzk6xssx4d/1) maintained at the Research Institute for Key Indicators and the Pennsylvania State University.

2. Methods

Alternate methodologies, logic models, and algorithms were developed directly from the Theory of Regulatory Compliance once it was determined that substantial regulatory compliance produced better results than full regulatory compliance. These methodologies created a differential monitoring or targeted monitoring approach based on risk assessment, which measures client morbidity and/or mortality when individual rule

*Corresponding author: Richard J. Fiene, Email: rjf8@psu.edu, Phone: 717-598-8908, ORCID iD: http://ORCID: 0000-0001-6095-5085.

[1] The term "facilities" is used when referring to programs and/or facilities.

[2] The term "rules" is used when referring to rules and/or regulations.

non-compliance is assessed, and the determination of key statistical predictors for overall regulatory compliance [3].

Briefly, the above methodologies provide cost-effective and efficient means for the ongoing monitoring of human service delivery systems by selecting and reviewing only those rules that either have a positive impact on clients, statistically predict overall regulatory compliance, or protect the health and safety of clients [3]. Based on regulatory compliance historical data, decisions could be made as to the frequency and depth of the reviews or inspections. Abbreviated reviews (inspections in which a subset of rules are measured), such as licensing key indicator rules or risk assessment rules, would only be done in those facilities having a history of high regulatory compliance. Those facilities with a history of high regulatory non-compliance would continue to receive full regulatory compliance reviews as they did in the past.

3. Results

Prior to 1979, it was always assumed that there was a linear relationship between regulatory compliance measures and program quality measures of human service facilities. In a study conducted in that year, which compared results from early care and education programs, in particular child care centers, this assumption did hold up when one went from low regulatory compliance to substantial regulatory compliance. However, the results from substantial regulatory compliance to full (100 percent) regulatory compliance did not show the same linear relationship. Rather, it showed that those programs that were in substantial instead of full compliance were actually scoring higher on the program quality measures.

Since 1979, this result has been replicated in many other early care and education delivery system studies, both nationally in the United States (Head Start) [1] and in several states (Georgia, Indiana, Pennsylvania) [2]. In all these studies, one finds a non-linear - rather than a linear - relationship between regulatory compliance and the overall quality of the facilities being assessed.

4. Discussion

Based on the results above, there are several assumptions within regulatory compliance that need to be reconsidered:

1. Public policies that require full (100 percent) compliance with all rules may not be in the best interest of the clients being served, nor an effective use of limited regulatory resources. Potentially, emphasis on substantial regulatory compliance may be a more effective and efficient public policy related to client outcomes when it comes to their health, safety, and quality of life. Note that substantial compliance is still very high regulatory compliance (99-97 percent compliance with all rules) and produces positive client outcomes. As stated above, regulatory compliance data are extremely skewed and not normally distributed. There is very little variance in the data and the majority of programs are in either full or substantial regulatory compliance.

2. If a jurisdiction focuses on a substantial regulatory compliance public policy it opens up many system enhancements, such as differential or targeted monitoring, risk assessment analysis, and statistical key indicator rules that have been demonstrated to be cost effective and efficient approaches to reviewing program performance. In a full regulatory compliance public policy approach, none of these system enhancements can be employed, with the possible exception of the key indicator approach as delineated in number four below.

3. If a jurisdiction takes the position that all rules are not equal, then a risk assessment or weighting approach becomes an alternative based on the assumption that certain rules place clients at greater risk of death, serious injury, or other types of harm.

4. Even if a jurisdiction does not have a licensing law that allows issuing licenses on the basis of substantial compliance, there is the possibility that key indicators could still be used for abbreviated reviews or inspections, if there is no prohibition in statute or regulation that expressly forbids the use of this approach, since key indicators statistically predict full regulatory compliance. In other words, all rules are statistically predicted to be in regulatory compliance based on the results of the key indicators. Therefore, technically, all rules have been reviewed albeit short of a full review or inspection.

5. Based on previous research, utilizing a risk assessment approach along with a key indicator approach is the most cost effective and efficient differential monitoring system model. The reason is that both predictive rules and those rules that place clients at greatest risk are always assessed when a site visit review or inspection is done. Many more jurisdictions use a risk assessment approach at this point, but there is a loss of predictive regulatory compliance by just using it.

6. Based on previous regulatory compliance history, only those facilities in high regulatory compliance would be eligible for abbreviated key indicator and risk assessment reviews, whereas those with a history of high regulatory non-compliance would continue to receive full regulatory compliance reviews. This gets at the essence of the differential monitoring approach, which is cost neutral. Regulatory resources may then be re-allocated from the abbreviated reviews to more in-depth full regulatory compliance reviews.

7. Based on the use of the key indicator and risk assessment methodologies within a differential monitoring approach, it is possible to identify over multiple jurisdictions if there are generic rules that meet the criteria of risk abatement and prediction. Such an application has occurred in the United States with the creation of early care

and education standards entitled *Caring for Our Children Basics*, published by the Administration for Children and Families, US Department of Health and Human Services (2015).

5. Conclusion

Regulatory compliance is relatively new in applying empirical evidence and basic scientific principles to its decision making. In the past, it had been dominated by case studies and long narrative reports that did not lend themselves to quantitative analysis. There is a need to more clearly apply empirical evidence and the scientific method to rule development. Certain assumptions, such as full regulatory compliance as a sound public policy, are lacking in empirical evidence. This treatise on a theory of regulatory compliance is provided for its heuristic value for both regulatory researchers and policymakers in rethinking some basic regulatory compliance assumptions. It is not about more or less, rules but finding the "right rules" that protect clients, predict overall regulatory compliance, and produce positive client outcomes.

6. Declaration of Conflicting Interest

The authors declare no conflicts of interest.

7. Article Information

This article was received March 14, 2019, in revised form April 11, 2019, and made available online May 9, 2019.

8. References

[1] Fiene, R. (2013). *Office of Head Start Key Indicator Project Report.* Retrieved from https://drfiene.files.wordpress.com/2018/04/rheadstart-report.pdf

[2] Fiene, R. (2014). *Georgia Child Care Licensing Study: Validating the Core Rule Differential Monitoring System.* Retrieved from https://drfiene.files.wordpress.com/2018/04/2014-georgia.pdf

[3] Fiene, R. (2016). *Early Childhood Program Quality Indicator and Improvement Model (ECPQIM) and Differential Monitoring Logic Model and Algorithm (DMLMA) Readings*. Retrieved from https://drfiene.files.wordpress.com/2018/03/1ecpqims-book-of-readings.pdf

International Journal of Child Care and Education Policy
2013, Vol. 7, No. 1, 1-15

A Comparison of International Child Care and US Child Care Using the Child Care Aware – NACCRRA (National Association of Child Care Resource and Referral Agencies) Child Care Benchmarks

Richard Fiene, Ph.D.

Affiliated Faculty
Prevention Research Center
The Pennsylvania State University

This is a first of its kind study comparing the USA to other world countries utilizing the Child Care Aware – NACCRRA Child Care Benchmarks related to health and safety rules and regulations. A team of researchers analyzed the child care/early care & education rules and regulations from the USA and a selected group of countries to do a comparative analysis using the Child Care Aware – NACCRRA benchmarking scoring protocol. The results from the analyses were somewhat unexpected in that the scores between the USA and the other countries were not as statistically significant in the overall scores. However, when more specific benchmarks were compared statistically significant differences did appear in the health & safety and professional development areas.

Key words: Child Care Quality, Comparisons of USA and International Child Care, Child Care Regulations.

Introduction

The purpose of this paper is to compare several countries (N =20) and the United States on the Child Care Aware – formerly NACCRRA (National Association of Child Care Resource and Referral Agencies) Child Care Benchmarks that have used extensively in the USA to compare state regulatory and monitoring policy and implementation. The use of these benchmarks has been very useful in comparing states in the USA on an agreed upon series of child care benchmarks that have a great deal of support in the research literature (AAP/APHA, 2012, 2013; NACCRRA 2007, 2009, 2011). Previous research (OCED, 2006) has focused on early care and education policies in other countries which was a very important

41 Grandview Avenue
Middletown, Pennsylvania USA 17057
717-944-5868 Phone and Fax
Fiene@psu.edu
March 2013 (revised and resubmitted July 2013)

first step in making comparisons across countries. This paper will expand upon this comparison in order to begin applying the NACCRRA benchmarks to other countries and establish a baseline between the USA and other countries related to regulatory review and analysis. This study is important because it provides a common rubric for making comparisons between the USA and other countries that is reliable and valid (NACCRRA 2007, 2009, 2011) related to regulatory analysis. As far as the author can determine from his extensive review of the literature, similar studies of this type have not been attempted utilizing a standardized rubric created by a major national child care organization. There have been other studies completed in which comparisons were made of other countries, the OCED (2006) Starting Strong II study and report is an excellent example of this type of

DIFFERENTIAL MONITORING LOGIC MODEL & ALGORITHM (DMLMA©) (Fiene, 2012): A 4th Generation ECPQIM – Early Childhood Program Quality Indicator Model

CI x PQ => RA + KI => DM

Definitions of Key Elements:

CI = Comprehensive Licensing Tool (Health and Safety)(***Caring for Our Children***)
PQ = ***ECERS-R, FDCRS-R, CLASS, CDPES*** (Caregiver/Child Interactions/Classroom Environment)
RA = Risk Assessment, (High Risk Rules)(***Stepping Stones***)
KI = Key Indicators (Predictor Rules)*(13 Key Indicators of Quality Child Care)*(NACCRRA Benchmarks)
DM = Differential Monitoring (How often to visit and what to review)

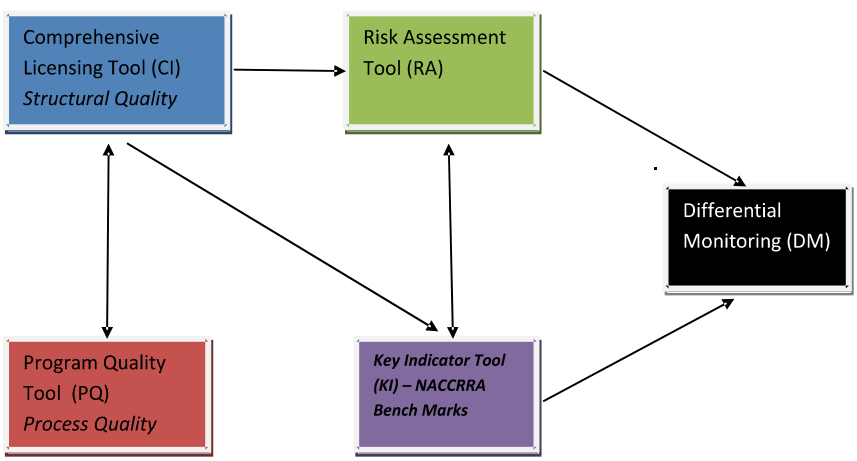


Figure 1.

analysis and is recommended reading for anyone interested in reviewing public policy analyses.

The child care benchmarks[1] utilized in this study are based upon the following key indicators: prevention of child abuse, immunizations, staff child ratio, group size, staff qualifications and training, supervision/discipline, fire drills, medication administration, emergency plan/contact, outdoor playground, inaccessibility of toxic substances, and proper hand washing/ diapering (NACCRRA 2007, 2009, 2011). These benchmarks are more based upon the structural aspects of quality rather than on the process aspects of quality. This is an important distinction between the USA approach and the other countries approaches that becomes important in the explanation of results later in this paper.

This paper also supports and expands the development of an Early Childhood Program Quality Indicator Model (ECPQIM)(Fiene & Nixon, 1985) which is in a 4th generation (Fiene, 2013) as a differential monitoring logic model & algorithm helping to guide the program monitoring of child care/early care & education programs (see Figure 1).

Method

Data Collection Process

Data collection was done on a 100 point scale which is delineated in Appendix 1 as developed by the Child Care Aware - NACCRRA Research Team. The same scoring protocol that was utilized in developing the 2007, 2009, and 2011 Reports and comparisons of states by Child Care Aware - NACCRRA was employed in this study in comparing the average scores of the states and the 20 countries. The 100 point scale consisted of 10 child care benchmarks each worth 10 points: ACR = Staff child ratios NAEYC Accreditation Standards met (R1); GS = Group size NAEYC Accreditation Standards met (R2); Director = Directors have bachelor's degree (R3); Teacher = Lead teacher has CDA or Associate degree (R4); Pre = Initial orientation training (R5); Inservice = 24 hours of ongoing training (R6); Clearance = Background check (R7); Devel = Six developmental domains (R8); Health = Health and safety recommendations (R9); and Parents = Parent Involvement (R10).

Data Scoring

The scoring protocol employed a total raw score approach of 100 points that was used to compare the countries on the 10 child care benchmarks in the aggregate. The scoring protocol also employed a standardized scoring approach (0 to 2 points) on each of the 10 child care benchmarks utilizing the following scale: 0.0 = Does not meet the Child Care Aware – NACCRRA Benchmarks; 0.5 = Marginally meets the Child Care Aware – NACCRRA Benchmarks; 1.0 = Partially meets the Child Care Aware – NACCRRA Benchmarks; 1.5 = Substantially meets the Child Care Aware – NACCRRA

Benchmarks; 2.0 = Fully meets the Child Care Aware – NACCRRA Benchmarks.

Data Collectors

A team of undergraduate and graduate research assistants[2] at the Pennsylvania State University were the data collectors in which each of them reviewed the child care/early childhood rules/regulations/standards from a specific country and scored the rules/regulations/standards on the Child Care Aware – NACCRRA 100 point raw score protocol and the standardized (0 – 2) scoring approach.

Data Sources

The child care regulations selected were for preschool age children only in child care center setting in the 20 countries. Geographically the governmental jurisdiction closest to the national capital was used if applicable national regulations could not be found. More than the final 20 countries selected were reviewed but several countries needed to be dropped because they did not meet the above criteria or the regulations could not be found in English. This was more a convenience sample rather than a stratified scientific sample, a limitation of this study.

Results

The results from this study and analysis were totally unexpected. The results indicated no statistically significant differences between the USA and the other countries selected (Australia, Belgium, Norway, Finland, Sweden, Ireland, United Kingdom, Italy, France, New Zealand, Mexico, Greece, Canada, Austria, Portugal, Philippines, Turkey, Pakistan, Nigeria, Denmark, and Spain – these countries were selected because of their availability of child care/early care & education rules and regulations as described previously above in Data Sources) when comparing the total scores on the 100 point scale; the USA average for all 50 states scored 58 while the 20 countries average score was 56. However, a very different scenario occurs when looking at the ten individual child care benchmarks using the standardized 0 – 2 scoring protocol. The 20 countries selected in this study scored statistically higher on the following child care benchmarks: Director ($t = 7.100$; $p < .0001$) and Teacher ($t = 7.632$; $p < .0001$) qualifications. The USA scored statistically higher on the following child care benchmarks: Health/Safety ($t = 6.157$; $p < .0001$), Staff Clearances ($t = 3.705$; $p < .01$), and Pre-Service ($t = 4.989$; $p < .001$) /In-Service training ($t = 2.534$; $p < .02$) (See Table 1 & Figure 2).

The results showed that both the USA and all other countries mean scores were 58 and 56 respectively on the 100 point scale – this is a raw scale score and not the standardized score (0 – 2 – see Table 1 and Figure 2) which was used in the comparisons for each benchmark. This is not a particularly good score if you think in terms of exams, but for states and countries with

Table 1

Mean Comparisons between USA and Twenty Countries on Child Care Aware – NACCRRA Benchmarks

Benchmark	Countries	USA	Significance
ACR (R1)	1.122	0.8462	not significant
GS (R2)	0.4063	0.5865	not significant
Director (R3)	**1.5625**	0.5	t = 7.100; p < .0001
Teacher (R4)	**1.6563**	0.4038	t = 7.632; p < .0001
Preservice (R5)	0.9375	**1.6731**	t = 4.989; p < .001
Inservice (R6)	0.6563	**1.0481**	t = 2.534; p < .02
Clearances (R7)	0.6094	**1.2404**	t = 3.705; p < .01
Development (R8)	1.6406	1.4519	not significant
Health(R9)	0.9844	**1.7404**	t = 6.157; p < .0001
Parent(R10)	1.5000	1.5385	not significant

Legend:
Child Care Aware - NACCRRA Benchmarks:
Parent = Parent Involvement (R10)
Health = Health and safety recommendations (R9)
Development = Six developmental domains (R8)
Clearances = Background check (R7)
Inservice = 24 hours of ongoing training (R6)
Preservice = Initial orientation training (R5)
Teacher = Lead teacher has CDA or Associate degree (R4)
Director = Directors have bachelor's degree (R3)
GS = Group size NAEYC Accreditation Standards met (R2)
ACR = Staff child ratios NAEYC Accreditation Standards met (R1)

Scoring:
0.0 = Does not meet Child Care Aware – NACCRRA Benchmarks.
0.5 = Marginally meets Child Care Aware – NACCRRA Benchmarks.
1.0 = Partially meets Child Care Aware – NACCRRA Benchmarks.
1.5 = Substantially meets Child Care Aware – NACCRRA Benchmarks.
2.0 = Fully meets Child Care Aware – NACCRRA Benchmarks.

vastly complex bureaucracies maybe this isn't as bad as it looks. Could it be that the USA is better than we think or is it that the USA and all other countries are providing just mediocre child care?!

The reason for using aggregate data in this study was to be consistent in how data have been collected in the USA utilizing the Child Care Aware – NACCRRA Scoring Protocol. This did delimit the potential analyses for this study and the recommendation would be made in future studies to unbundle

the results so that more detailed comparisons could be made. As mentioned in the introduction, the purpose of this study was to provide an initial baseline comparison between the USA and other countries on the Child Care Aware – NACCRRA Scoring Protocol.

Discussion

The purpose of this study was to extend the Child Care Aware - NACCRRA Child Care Benchmarks Scoring Protocol to an international sample comparison. As has been done by the National Science Foundation with math and science testing, these same types of comparisons have been made with the USA not fairing all that well on the math and science comparisons.

It appears that when it comes to child care benchmarks the USA actually appears to be in better shape than many advocates and experts would have thought when compared to other countries or is it that the other countries are providing the same form of mediocre care as it relates to these child care benchmarks. Remember that these benchmarks are heavily weighted towards the structural side of quality

Legend:
Child Care Aware - NACCRRA Benchmarks:
Parents = Parent Involvement (R10)
Health = Health and safety recommendations (R9)
Devel = Six developmental domains (R8)
Clearance = Background check (R7)
Inservice = 24 hours of ongoing training (R6)
Pre = Initial orientation training (R5)
Teacher = Lead teacher has CDA or Associate degree (R4)
Director = Directors have bachelor's degree (R3)
GS = Group size NAEYC Accreditation Standards met (R2)
ACR = Staff child ratios NAEYC Accreditation Standards met (R1)

Scoring:
0.0 = Does not meet Child Care Aware – NACCRRA Benchmarks.
0.5 = Marginally meets Child Care Aware – NACCRRA Benchmarks.
1.0 = Partially meets Child Care Aware – NACCRRA Benchmarks.
1.5 = Substantially meets Child Care Aware – NACCRRA Benchmarks.
2.0 = Fully meets Child Care Aware – NACCRRA Benchmarks.

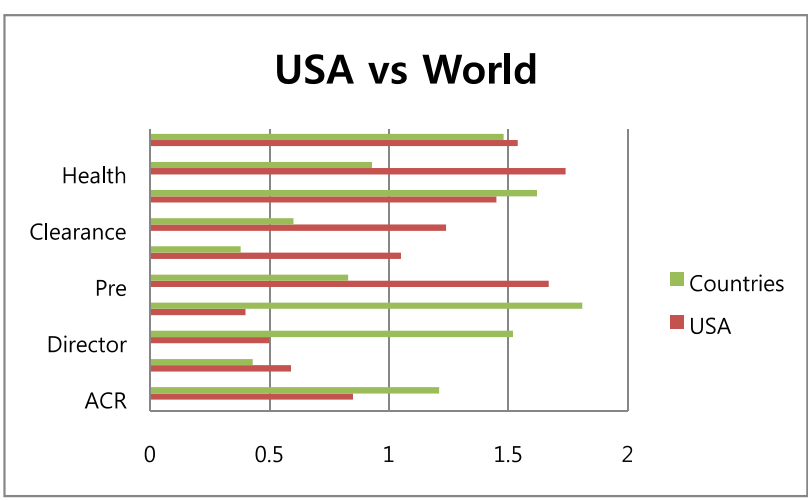


Figure 1. *Mean Comparisons between USA and Twenty Countries on Child Care Aware – NACCRRA Benchmarks*

rather than the process side of quality.

However, when the individual benchmarks are analyzed then certain patterns occur which seem very consistent with the previous research literature. The 20 countries scored higher on the staffing benchmarks while the USA scored higher on the training and health/safety benchmarks. Clearly this is an indication reflecting public policy in the other countries as versus the USA. Many other countries place more emphasis on the process aspects of quality which involve staff and staff interactions with children. The USA has focused more on the structural aspects of quality which involve health & safety especially in the state licensing of child care. These structural aspects of quality are more easily quantifiable in state rules and regulations which is the locus of control for the licensing of child care. Since the USA does not have national standards that are required (the USA does have national health and safety standards that are recommended practice, such as Caring for Our Children (2012)) as is the case in so many of the countries in this study, this may provide a possible explanation for the results of this study. It will be interesting to see how Quality Rating and Improvement Systems (QRIS) which usually have some process standards impact this overall balance of structural and process aspects of quality. This is an area that needs additional research and more in-depth analysis.

So what does this tell us. I think it is a warning call as has been put forth by Child Care Aware - NACCRRA that we still have a lot of additional work to do in improving child care, not only in the USA, but worldwide. Just as the Child Care Aware -NACCRRA Report Cards (2007, 2009, 2011) have played a role in making positive change in the child care benchmarks over time; we need to expand this reporting and change to a world wide focus. There is clearly the need to expand from the present analysis of 20 countries and the USA to other countries throughout the world and to track changes over time as Child Care Aware/NACCRRA has done.

Another area of concern within the USA and I am sure in other countries as economies have begun their slow recovery from the economic downturn of 2008 – 2010 is to do more with less. One such approach being explored in the USA is called differential monitoring which helps to re-allocate limited resources in a more cost effective and efficient manner via a risk assessment and key indicator approach. I hope that this comparison utilizing the Child Care Aware – NACCRRA Benchmarking Scoring Protocol and introducing the Early Childhood Program Quality Indicator Model/Differential Monitoring Logic Model and Algorithm (Fiene, 2013) within an international context as first steps in making that happen.

References

AAP/APHA (2012). *Caring for our children*, Washington, D.C.: American Public Health Association.

AAP/APHA (2013). *Stepping Stones*, Washington, D.C.: American Public Health Association.

Child Care Award - NACCRRA, (2007). *We Can Do Better: NACCRRA's Ranking of State Child Care Center Standards and Oversight*, Washington, D.C.: National Association of Child Care Resource and Referral Agencies.

Child Care Aware - NACCRRA, (2009). *We Can Do Better: 2009 Update: NACCRRA's Ranking of State Child Care Center Regulations and Oversight*, Washington, D.C.: National Association of Child Care Resource and Referral Agencies.

Child Care Aware - NACCRRA, (2011). *Leaving Children to Chance: NACCRRA's Ranking of State Standards and Oversight in Child Care Centers*, Washington, D.C.: National Association of Child Care Resource and Referral Agencies

Fiene, R. (2013). *Differential monitoring logic model and algorithm*. Middletown, Pennsylvania: Research Institute for Key Indicators.

Fiene, R. & Nixon, M. (1985). Instrument based program monitoring and the indicator checklist for child care, *Child Care Quarterly, 14*(3), 198-214.

OCED (2006). *Starting strong II*. Paris, France: Organization for Economic Co-Operation and Development Publishing.

Notes

[1] In the licensing literature these child care benchmarks are usually referred to as key indicators (Fiene, 2013). Please see Figure 1 which delineates where within a program monitoring system these benchmarks would appear and could be utilized.

[2] The following individuals played key data collection roles as research assistants in the compilation of this study: Melissa Cave, Ashley Le, Breanna Green, Corrie Podschlne, Sherrie Laporta, Ashley Edwards, Laura Hartranft, Gissell Reyes, Janet Lazur, Kayma Freeman, Jessica White, Karen Mapp, and Lindsay Bitler.

Appendix 1

Benchmark criteria for *We Can Do Better: NACCRRA Ranking of State Child Care Center Regulations:2011 Update* were developed by Child Care Aware - NACCRRA and have been used for the 2007, 2009 and 2011 We Can Do Better reports. The rationale for each standard, including research evidence of its importance in quality care, is noted in each section of the report and in previous reports. Each of the 10 regulation benchmarks were scored with a value ranging from one to 10 points, depending on how closely the state met the benchmark, for a maximum total of 100 points. In cases where states permit several different options for complying (e.g., complying with director or teacher qualifications), the minimum allowed was used. This information was used to generate state sheets with scores for each standard.

Scoring Methods for NACCRRA Ranking of State Child Care Center Regulations (R)

Question	Scoring method
Regulation 1. Staff:child ratio requirements comply with NAEYC accreditation standards.	

6 mo	9 mo	18 mo	27 mo	3 yr	4 yr	5 yr
1:4	*1:4*	*1:4*	*1:4*	*1:9*	*1:10*	*1:10*

Number of ratios in compliance with NAEYC standards	Score
7 ratios	10
6 ratios	9
5 ratios	8
4 ratios	7
3 ratios	5
2 ratios	3
1 ratios	1

R2. Group size requirements are in compliance with NAEYC accreditation standards.

6 mo	9 mo	18 mo	27 mo	3 yr	4 yr	5 yr
8	*8*	*8*	*8*	*18*	*20*	*20*

Number of group sizes in compliance with NAEYC standards	Score
7 ratios	10
6 ratios	9
5 ratios	8
4 ratios	7
3 ratios	5
2 ratios	3
1 ratios	1

R3. Center directors are required to have a bachelor's degree of higher in early childhood education or a related field.

Director education requirement	Score
Bachelor's degree in any field	10
College directors certification	7
Any associate degree	5
CDA	5
Clock hours/less than associate degree	2
High school or less	0

R4. Lead teachers are required to have a Child Development Associate (CDA) credential or an associate degree in early childhood education or related field.

Lead teacher education requirement	Score
CDA/associate degree or better	10
State Credential	5
Clock Hours in ECE	2
High School/GED	2
Less than High School	0

R5. Lead teachers are required to have initial training, including:

- Orientation.
- Fire safety.
- Other health and safety issues.
- At least one staff member certified in first aid must be present when children are in care.
- At least one staff member who is certified in CPR must be present when children are in care.

Number of areas training is required	Score
Five areas	10
Four areas	8
Three areas	6
Two areas	4
One area	2
None	0

R6. Lead teachers are required to have 24 hours or more of annual training.

Ongoing training ≥	Score
24 Hours	10
18 hours	7
12 hours	5
6 hours	2
None	0

R7. A comprehensive background check is required for child care providers.

- Use of fingerprints to check state records.
- Check FBI records.
- Check state child abuse registry
- Check sex offender registry.
- Criminal history check.

Number of Background checks completed	Score
Five checks	10
Four checks	8
Three checks	6
Two checks	4
One check	2
None	0

R8. Child care centers are required to offer program activities that address all six child development domains

- Language/literacy.
- Cognitive.
- Social.
- Emotional.
- Physical.
- Cultural.

Developmental domains addressed	Score
6 domains	10
5 domains	9
4 domains	7
3 domains	5
2 domains	3
1 domain	1
None	0

R9. Child care centers are required to follow 10 recommended health and safety practices.

- Immunizations.
- Guidance/discipline.
- Diapering and handwashing.
- Fire drills.
- Medication administration.
- SIDS prevention.
- Emergency preparedness.
- Playground surfaces.
- Hazardous materials.
- Incidence reporting.

Standards addressed	Score	Standards addressed	Score
10	10	5	5
9	9	4	4
8	8	3	3
7	7	2	2
6	6	1	1
Allowing corporal punishment is an automatic zero			

R10. Child care centers are required to:

- Encourage parent involvement.
- Require daily or ongoing communication with parents.
- Allow parental access any time their children are in care.

Number of items required	Score
Three items	10
Two items	7
One item	3
None	0

Appendix 2

These were the countries included in these analyses: Australia, Belgium, Norway, Finland, Sweden, Ireland, United Kingdom, Italy, France, New Zealand, Mexico, Greece, Canada, Austria, Portugal, Philippines, Turkey, Pakistan, Nigeria, Denmark, Spain, and the USA which included all 50 states.

big data and cognitive computing

MDPI

The Intersection of Structural Quality and Process Quality in Building an Integrated Program Monitoring Systems Approach Using Artificial Intelligence

Richard Fiene PhD 0000-0001-6095-5085

Research Institute for Key Indicators Data Laboratory

Penn State Edna Bennett Pierce Prevention Research Center

October 2025

Abstract

The Child Care and Early Education (CCEE) Heart Monitor (CCEEHM) is introduced as a new Integrated Program Monitoring System's Approach to assessing both structural and process quality in one platform. It builds upon the Contact Hour (CH) metric and the Key Indicator Methodology (KIM) that have been introduced in the CCEE licensing and monitoring field. The CCEEHM expands the use of the CH and KIM methods by integrating key elements from both structural and process quality into a software application utilizing artificial intelligence that can be used by staff, licensors, and quality assessors. The CCEEHM draws indicators from licensing, regulatory compliance, quality rating and improvement systems, and other quality initiatives, such as accreditation, and professional development and technical assistance systems.

Key Words: Artificial Intelligence, Big Data, Structural Quality, Process Quality, Contact Hour Metric, Regulatory Compliance, Key Indicator Methodology, Integrated Program Monitoring

Introduction

The Child Care and Early Education (CCEE) field needs a means to monitor the key elements of structural and process quality in a unified framework. The theory of regulatory compliance has been suggested as this unifying framework for structural and process quality (Fiene, 2019; 2021; 2025a,b); but at a more practical level what could be used to essentially unify the monitoring and measurement of both structural and process quality. Generally, structural and process quality are measured separately from each other by using very separate and distinct tools utilized by licensing inspectors and quality observers (Kontos & Fiene, 1987). This research paper will build off several measurement concepts (binary and ordinal measurement)

that deal with the creation of a new Contact Hour (CH) metric replacing measuring compliance with adult-child ratios while unifying structural quality with process quality. With this new unification of structural and process quality, it will help to build a more Integrated Monitoring Systems Approach (Freer & Fiene, 2023) which should go a long way in complementing the measurement strategies employed in licensing and quality rating and improvement systems that have proliferated in the child care and early education field.

Let's begin by placing some context on the title of this new Child Care and Early Education Heart Monitor. What do we mean by heart monitor? Within the research literature in determining the levels of quality generally these levels are broken into two distinctive categories, those that deal with structural quality, such as staff child ratios, group size, etc. Essentially health and safety or licensing rules and regulations. The interactions amongst the staff and children generally fall under the process quality side of the equation. But this is really the "heart" of quality. This is where the magic occurs, the so-called "dance" between the adult and the child(ren). All the structural quality rules and regulations are important in protecting children and keeping them healthy but the interaction of child and adult is where the action occurs. So what is being proposed is to combine these two categories of quality together into one system, placing the measurement and the monitoring of process quality squarely within the structural measurement strategy, the Contact Hour (CH) metric. Another way of looking at this relationship is by combining the two pillars of regulatory compliance "Do no harm" and "Do good" into a unified single platform where they build upon each other.

This framework will be developed within this paper by fully describing the Contact Hour metric (Fiene & Stevens, 2021) and a newly created CCEE Quality Indicator tool (Fiene, 2024) that will measure the quality enhancements within the Contact Hour metric and do this within an App (software application) that can be downloaded and it will produce the scores based upon reviewing specific documents and observations within a child care and early education program. This new Child Care and Early Education Heart Monitor

(CCEEHM) should be both cost effective and efficient being based upon the key indicator methodology (Fiene & Nixon, 1985) and having it developed into an App (software application) should make it particularly easy to use for licensors, assessors, or observers since all the scoring would be done by the CCEEHM App.

Let's continue by delving into the Contact Hour (CH) metric (Fiene & Stevens, 2021). The Contact Hour metric has been proposed as a more effective and efficient metric for measuring compliance with adult-child ratios and group sizes in CCEE programs, and for monitoring the spread of infectious diseases. It is simple to apply by just asking 6 questions about when children arrive and leave a CCEE program and how many staff are present in a particular classroom (See the Methodology section for the questions and algorithms). Once that is done a trapezoidal model is built in which compliance with staff child and group size rules can be determined. Regulatory compliance is determined by comparing the resultant area to an ideal level of contact between staff and children. This Introductory section is followed by the tool that would be used for determining the Contact Hour metric as well as the Program Quality Indicators (PQI) that need to be measured in the Methodology section. Also, there is the Scoring Protocol to be used in determining the level of quality and a screen shot of the opening page of the CCEEHM App that has been designed to measure compliance with the tools for CH and PQI in the Results section. The Discussion section is provided in which a hypothetical example of two programs, one of high quality and one of low quality, are delineated demonstrating the scoring protocol in greater detail.

In determining the results, the Contact Hours (CH) are dealt with as absolute values but let's enhance this result by moving it from an absolute value to one that is more relative by introducing process quality measures such as the Program Quality Indicators (PQI). The PQI portion of the tool has a good deal of observations that need to be made in classrooms. To do this, it would take 1000's of observations to fill the Contact Hour trapezoidal model which is not realistic. But let's let Artificial Intelligence (AI) do the observing and training of AI in what constitutes the various quality levels on the respective CH/PQI tool. By

using AI and having video cameras in each of the classrooms to be assessed, this becomes doable. The CH/PQI observer would be able to collect the data by observing and assessing what it sees via the video cameras installed in the classrooms. Summary measurements would be made on an hourly basis and recorded as part of the Contact Hour trapezoidal model. At the end of the day, there would be a relative value utilized in this model rather than the absolute value that has been used in the past to determine structural quality compliance with adult-child ratio and group size. For example, if a CCEE program classroom exceeded the area of the trapezoidal model it would be out of compliance and if it were within the area of the trapezoidal model it was in compliance (see the following Methodology section related to the calculation of the Contact Hour metric). By adding the PQI data, it changes this metric totally by adding process quality measures which can be measured on a 1-4 ordinal scale, similar to accreditation systems or an ordinal (1-7) scale, similar to many program quality tools, such as the Environmental Rating Scales.

This approach will get at the ***Heart of CCEE monitoring, "process quality"***, measuring the interactions amongst staff and children in an ongoing fashion. It moves the needle from being structural to process quality providing an intersection of both components of quality. The AI approach will also help to address the issues related to bias in regulatory compliance observing and decision making by inspectors/observers. By training the AI PQI Observers there should be greater certainty established in making the right decisions related to specific quality elements (Fiene, 2025c). Just as in establishing inter-rater reliability with human observers, the same can be done with the PQI AI Observers but there will be less drift with AI.

The next section describes the Contact Hour Metric methodology in detail and provides the Program Quality Indicators (PQI) that are part of the CCEEHM App. This methodology provides the meat of the new Integrated Program Monitoring Systems Approach. In fact a human observer could use these two sections and then manually use the CCEEHM App for doing their data entry. The App would then do all the scoring for the individual assessor (See the Results Section which contains the opening screen to the App as well as the scoring protocol).

Methodology: Contact Hour (CH) Metric - The Structural Quality Component

One starts the Contact Hour (CH) metric methodology by asking the following six questions (The six questions should be asked of each grouping that is defined by a classroom or a well-defined group within each classroom tied to a specific adult-child ratio.):

1.When does your first teaching staff arrive or when does your facility open (TO1)?
2.When does your last teaching staff leave or when does your facility close (TO2)?
3.Number of teaching/caregiving staff (TA)?
4.Number of children on your maximum enrollment day (NC)?
5.When does your last child arrive (TH1)?
6.When does your first child leave (TH2)?

After getting the answers to these questions, the following formulae can be used to determine contact hours (CH) based upon the relationship between when the children arrive and leave (TH) and how long the facility is open (TO):

$$CH = ((NC\ (TO + TH)) / 2) / TA;$$

$$CH = (NC \times TO) / TA;$$

$$CH = ((NC \times TO) / 2) / TA;$$

$$CH = (NC^2) / TA$$

Where: CH = Contact Hours; NC = Number of Children; TO = Total number of hours the facility is open (TO2 - TO1); TA = Total number of teaching staff, and TH = Total number of hours at full enrollment (TH2 - TH1).

By knowing the number of contact hours (CH) it will be possible to rank order the exposure time of adults with children. Theoretically, this metric could then be used to determine that the greater contact hours is correlated with the increased non-regulatory compliance with adult-child ratios as determined in the below table (Table 1).

Table 1: Contact Hour (CH) Conversion Table (RS Model(1.0)) (Fiene, 2020©)

Taking into Account Exposure Time and Density

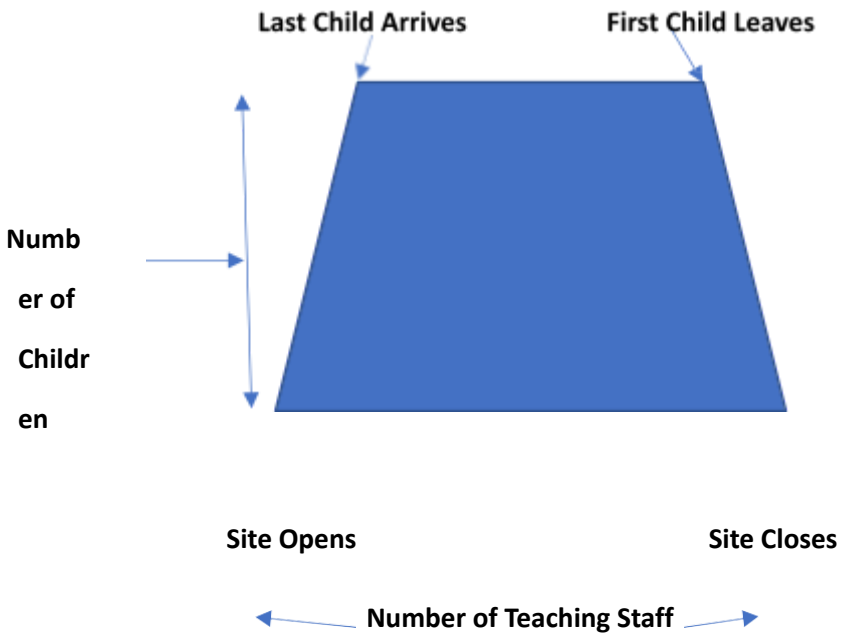
Group Size, Staff Child Ratio, Number of Children and Staff

<---------- Adult-Child Ratios (Relatively Weighted Contact Hours)---------->

NC	CH	1:1	2:1	3:1	4:1	5:1	6:1	7:1	8:1	9:1	10:1	11:1	12:1	13:1	14:1	15:1
1	8	**8**	8	8	8	8	8	8	8	8	8	8	8	8	8	8
2	16	**8**	**16**	16	16	16	16	16	16	16	16	16	16	16	16	16
3	24	**8**	12	**24**	24	24	24	24	24	24	24	24	24	24	24	24
4	32	**8**	**16**	16	**32**	32	32	32	32	32	32	32	32	32	32	32
5	40	**8**	13	20	20	**40**	40	40	40	40	40	40	40	40	40	40
6	48	**8**	**16**	**24**	24	24	**48**	48	48	48	48	48	48	48	48	48
7	56	**8**	14	19	28	28	28	**56**	56	56	56	56	56	56	56	56
8	64	**8**	**16**	21	**32**	32	32	32	**64**	64	64	64	64	64	64	64
9	72	**8**	14	**24**	24	36	36	36	36	**72**	72	72	72	72	72	72
10	80	**8**	**16**	20	27	**40**	40	40	40	40	**80**	80	80	80	80	80
11	88	**8**	15	22	29	29	44	44	44	44	44	**88**	88	88	88	88
12	96	**8**	**16**	**24**	**32**	32	**48**	48	48	48	48	48	**96**	96	96	96
13	104	**8**	15	21	26	35	35	52	52	52	52	52	52	**104**	104	104
14	112	**8**	**16**	22	28	37	37	**56**	56	56	56	56	56	56	**112**	112
15	120	**8**	15	**24**	30	**40**	40	40	60	60	60	60	60	60	60	**120**
16	128	**8**	**16**	21	**32**	32	43	43	**64**	64	64	64	64	64	64	64
17	136	**8**	15	23	27	34	45	45	45	68	68	68	68	68	68	68
18	144	**8**	**16**	**24**	29	36	**48**	48	48	**72**	72	72	72	72	72	72
19	152	**8**	15	22	30	38	38	51	51	51	76	76	76	76	76	76
20	160	**8**	**16**	23	**32**	**40**	40	53	53	53	**80**	80	80	80	80	80
21	168	**8**	15	**24**	28	34	42	**56**	56	56	56	84	84	84	84	84
22	176	**8**	**16**	22	29	35	44	44	59	59	59	**88**	88	88	88	88
23	184	**8**	15	23	31	37	46	46	61	61	61	61	92	92	92	92
24	192	**8**	**16**	**24**	**32**	38	**48**	48	**64**	64	64	64	**96**	96	96	96
25	200	**8**	15	22	29	**40**	40	50	50	67	67	67	67	100	100	100
26	208	**8**	**16**	23	30	35	42	52	52	69	69	69	69	**104**	104	104
27	216	**8**	15	**24**	31	36	43	54	54	**72**	72	72	72	72	108	108
28	224	**8**	**16**	22	**32**	37	45	**56**	56	56	75	75	75	75	**112**	112
29	232	**8**	15	23	29	39	46	46	58	58	77	77	77	77	77	116
30	240	**8**	**16**	**24**	30	**40**	**48**	48	60	60	80	80	80	80	80	**120**

This table is based upon the assumptions that the child care is 8 hours in length (TO) and that the full enrollment is present for the full 8 hours (TH). This is unlikely to ever occur but it gives us a reference point to measure adult child contact hours in the most efficient manner. Based upon the relationship between TO and TH based upon the algorithms, select from one of the formulae from the previous page (formulae 1 - 4) to determine how well the actual Relatively Weighted Contact Hours (RWCH) match with this table. If the RWCH exceed the respective RWCH in this table, then the facility would be over ratio on ACR standards, in other words, they would be overpopulated.

Figure 1: Contact Hour Diagram Paradigm and Schematic



The above diagram (Figure 1) depicts how the number of staff and children help to construct the contact hour formula. Depending on when the children arrive and leave could change the shape from a trapezoid to a rectangle or square or triangle. Please see the following potential density distributions which could impact these changes in the above contact hour diagram.

Potential Density Distributions
Taking into Account Number of Children, Staff, and Exposure Time

Here are some basic key relationships or elements related to the Contact Hour (CH) methodology.

- RWCH = ACR
- CH = GS = NC
- NC and CH are highly correlated
- ACR and GS are static, not dynamic
- CH makes them dynamic by making them 2-D by adding in Time (T)
- ΣACR = GS
- GS = total number of children NC
- ACR = children / adult

ACR = Adult Child Ratio, GS = Group Size, RWCH = Relatively Weighted Contact Hours, NC = Number of Children.

Possible Density Displays of Contact Hours (Horizontal Axis = Time (T); Vertical Axis = NC):

This density distribution should result in the lowest CH but probably not very likely to occur. Essentially what would happen is that full enrollment would be a single point which means that the last child arrives when the first child is leaving. Very unlikely but possible.

This density distribution is probably the most likely scenario when it comes to CH in which the children gradually, albeit rather steeply, arrive at the facility and also leave the facility gradually. They don't all show up at the same time nor leave at the same time. However, the arriving and leaving will be a rather close time frame.

This scenario is unlikely but is used as the reference point for CH because it provides the most efficient model. This is where all the children arrive and leave at the same time. Very unlikely, but I guess it could happen. The important element here is its efficiency in that all contact hours are covered, so although a lesser amount of CH is not as efficient it does demonstrate compliance with ACR and GS which is one of the purposes of CH. As the bottom two distributions will demonstrate, CHs above this level would either depict a program that is open for an extended time or where there are too many children present and the facility is out of compliance with GS and/or ACR.

This distribution would indicate that the facility is open for an extended time and exceeds the number of total CH as depicted in the reference square standard. Although not out of compliance with GS or ACR, this could become a determining factor when looking at the potential overall exposure of adults and children when we are concerned about the spread of an infectious diseases, such as what happened with COVID19. Are facilities that are high on a CH measurement more prone to the spread of infectious diseases?

This depiction clearly indicates a very high CH and non-compliance with ACR and GS. This is the reason for designing the CH methodology which was to determine these levels of regulatory compliance as its focus.

Program Quality Indicators (PQI) - The Process Quality Component

This section provides the program quality indicators (PQI) which along with the previous contact hour metric dealing with staff child ratios and group sizes constitutes the new Integrated Program Monitoring system: CCEE Heart Monitor (CCEEHM App). These PQI were validated in a study in the province of Saskatchewan (Fiene, 2024).

The PQI represents staffing, program, parental involvement and key interactional observation indicators drawn from key indicator studies from 1980 - 2020 involving quality rating and improvement systems (QRIS), professional development, and program quality initiative observational studies. These indicators provide the process quality within the context of the structural quality provided by the contact hour metric depicted in the previous section. Both the contact hour and these PQI are intended to be used in an integrated fashion and compliance should be measured on both domains. By doing this a picture of structural and process quality will be possible.

By utilizing this new integrated program monitoring system it will provide a cost effective and efficient system for jurisdictions around the world. These metrics are based upon research studies completed in the USA and Canada from 2020-2024 (Fiene, 2025a,b,c).

INDICATOR 1): Number of ECE III Educators (AA and BA Level ECE Educators)

AI will review staff records to determine the number of staff who have these credentials in early childhood education. Record the number of ECEs with the appropriate qualifications and divide them by the total number of ECEs to come up with a percent for the center.

How to Measure:

Go to a Staff Information Summary form to obtain the data for this item. Under Certification, look for the following: Certification Date and Certification Level (Highest ECE Level Certified). The certification date should be earlier than the date of the review and the actual level of the certification. In this case, we are interested in the number of (ECEIII's). Record the number of ECEIII working at least 65 hours/month. Then record the number of total teaching staff working at least 65 hours/month below as well. Teaching staff is defined as staff who have a responsibility for working with the children and the programming. Determine the percentage by dividing the total number of staff into the total number of ECEIII Certified teaching staff, ECEIII Certified teaching staff is the numerator, and the total number of teaching staff is the denominator (ECEIII/Total number of teaching staff x 100% = Percent).

Scoring for PQI 1:

The total number of ECEIII Certified teaching staff _____

(1.1) The total number of teaching staff _____(1.2)

Total ECEIII teaching staff divided by the total number of teaching staff _____

(%). Then based on the percentage, you can find the score of 1-4 as per the chart below.

Circle the Appropriate Level	***1 = 0 to 25%***	***2= 26 to 50%***	***3 = 51 to 75%***	***4 = 76 to 100%***

INDICATOR 2): Stimulating and Dynamic Environment

The criteria for measuring this are drawn from Play and Exploration Guides that should be present in all CCEE programs. The program should be child centered. Children are viewed as competent learners, and they have the freedom to access classroom materials independently without adult intervention. The children are provided with meaningful choices through activity/learning centers. There is evidence of the children's interests and their projects in the learning environment.

How to Measure:

Below is the checklist of items that should be present to assess if the environment is both stimulating and dynamic for the children. You will want to observe that the following items are occurring in the classroom first. If you do not actually observe it occurring, then check the program plan to find documentation that it normally occurs but you just did not observe today. The checklist items would be found in *Play and Exploration* foundational materials.

Quality Early Learning Environments (Please record all that you observe Y or N):

1. Co-teaching is evident. Y/N _____(2.1)
2. Children are viewed as competent learners & can access materials independently. Y/N ____(2.2)
3. Authentic and meaningful materials are used with children. Y/N _____(2.3)
4. Children are provided with meaningful choices. Y/N _____(2.4)
5. Children's work, art and photos are displayed respectfully. Y/N _____(2.5)
6. Family photos are displayed in the early learning program. Y/N _____(2.6)
7. Documentation of learning is displayed and discusses holistic development. Y/N _____(2.7)
8. Environment reflects the culture and beliefs of the children, families and staff. Y/N _____(2.8)
9. Variety of books & other print materials are available throughout the classroom Y/N _____(2.9)
10. A variety of writing materials are accessible to children most of the time. Y/N _____(2.10)
11. There is evidence of the children's interests & projects in the classroom. Y/N _____(2.11)

Scoring for PQI 2:

Total up the number of items where you recorded a "Y" above that you observed (curriculum or in classrooms), divide by 11 x 100% to come up with a percent and record here _____%. Then based on the percentage, you can find the score of 1-4 as per the chart below.

Circle the Appropriate Level	***1 = 0 to 25%***	***2= 26 to 50%***	***3 = 51 to 75%***	***4 = 76 to 100%***

INDICATOR 3): Developmentally Appropriate Curriculum Based on Assessments of Each Child

The key for this quality key indicator is that the program is following an individualized prescribed planning document when it comes to curriculum. It does not mean it is a canned program, in fact, it shouldn't if it is based upon the individual needs of each child's developmental assessment. The assessor will ask to see what is used to guide the curriculum. There should be a written document that clearly delineates the parameters of the philosophy, activities, guidance, and resources needed for the particular curricular approach. There should also be a developmental assessment which is clearly tied to the curriculum. The developmental assessment can be home-grown or a more standardized off-the- shelf type of assessment, the key being its ability to inform the various aspects of the curriculum. The purpose of the assessments is not to compare children but rather to compare the developmental progress of individual children as they experience the activities of the curriculum.

The following key elements should be present when assessing this quality indicator.

- 1) The program practices emergent curriculum, allowing the interests of the children to determine the learning content. The curriculum is informed by individual developmental assessments of each child in the respective classrooms.
- 2) The children and educators are co-learners in the exploration of projects.
- 3) Learning activities of the children are documented, displayed in the learning environment and used to plan further learning activities. This can be assessed developmentally.

How to Measure:

Take a sample of 10 individual children's records and consider the above three elements for EACH record. You should be asking yourself if there is a clear link between an assessment and the developmentally appropriate curriculum so that an individualized learning approach is being undertaken and each child's developmental needs are taken into consideration. These records could be formal, such as portfolios kept for each child or a more informal, anecdotal type of record keeping. The key is that there is a record that can be looked at. It is not adequate if the teacher says they do it from memory – it needs to be written down and documented.

Cross check the child's record to the actual curriculum. Record all the instances (Y's) in which this occurs. All three blocks need to be checked for each record (1-10).

Emergent Curriculum is Practiced (3.1)

1 Y/N	2 Y/N	3 Y/N	4 Y/N	5 Y/N	6 Y/N	7 Y/N	8 Y/N	9 Y/N	10 Y/N

Key Element 1 +

Children and Educators are Co-learners (3.2)

1 Y/N	2 Y/N	3 Y/N	4 Y/N	5 Y/N	6 Y/N	7 Y/N	8 Y/N	9 Y/N	10 Y/N

Key Element 2 +

Learning Activities are Documented and Displayed and Used to Plan Future Learning (3.3)

1 Y/N	2 Y/N	3 Y/N	4 Y/N	5 Y/N	6 Y/N	7 Y/N	8 Y/N	9 Y/N	10 Y/N

Key Element 3 +

All three key elements must have a Y to get an overall score of Y. If all three key elements have a Y for that individual record, then record Y in the corresponding block in the overall score.

1 Ys =	2 Ys =	3 Ys =	4 Ys =	5 Ys =	6 Ys =	7 Ys =	8 Ys =	9 Ys =	10 Ys =

= Total of All Three Key Elements (3.4)

Scoring for PQI 3:

The number of positive records (all Ys for all three elements) where there is a crosswalk from developmental assessment to curriculum _____

Percent of positive records (all Ys) (divide the number of positive records by 10 x 100%) _____

%. Then based on the percentage, you can find the score of 1-4 as per the chart below.

Circle the Appropriate Level	***1 = 0 to 25%***	***2= 26 to 50%***	***3 = 51 to 75%***	***4 = 76 to 100%***

INDICATOR 4): Opportunities for Staff and Families to Get to Know Each Other

There should be activities both within the center as well as off site where staff and parents have opportunities to meet and greet each other. Communication with family members is documented and enables early childhood providers to assess the need for follow-up. Early childhood providers hold regular office hours when they are available to talk with family members either in person or by phone. Family members are encouraged to lead the conversation and to raise any questions or concerns.

How to Measure:

Look for the following 3 examples in policies developed by the program and determine if they have been carried out with families. It will be necessary to interview staff to complete this indicator if you do not find the three examples in policies:

1. The program provides communication, education, and informational materials & opportunities for families that are delivered in a way that meets their diverse needs. Y/N_ (4.1)
2. The program communicates with families using different modes of communication, and at least one mode promotes two-way communication. Y/N ____(4.2)
3. The program demonstrates respect and engages in ongoing two-way communication. The program respects each family's strengths, choices, & goals for their children. Y/N __ (4.3)

Scoring for PQI 4:

Record the number of Yes's (Y's): _____(Range: 0 – 3) (Divide by 3 x 100% = _____%). Then based on the percentage, you can find the score of 1-4 as per the chart below.

Circle the Appropriate Level	***1 = 0 to 25%***	***2= 26 to 50%***	***3 = 51 to 75%***	***4 = 76 to 100%***

INDICATOR 5): Families Receive Information on Their Child's Progress Regularly Using a

Formal Mechanism

Based upon Indicator #3 above, the information gleaned from the developmental assessments should be the focus of the report or parent conference. Parental feedback about the assessment and how it compares to their experiences at home would be an excellent comparison point. All these interactions should be done in a culturally and linguistically appropriate way representing the parents being served.

How to Measure:

Look for the following four examples in policies developed by the program and determine if they have been carried out with families. Record the number of reports completed or parent conferences over the past year. It will be necessary to interview staff to complete this indicator if you cannot determine from records that the conferences or reports were completed.

- 1) The program does have regularly scheduled (at least 2xs/year) parent conferences in which the children's developmental progress is discussed AND provides the family with a report of their child's developmental progress. Y/N _____(5.1) (Score 3 points). If "Yes" then go to Number 4. If "No", then go to numbers 2 and 3.
- 2) The program has regularly scheduled (at least 2xs/year) parent conferences in which the children's developmental progress is discussed, but it does not provide a report to the parents on their child's developmental progress. Y/N ___(5.2) (Score 2 points).
- 3) If the program does not have regularly scheduled (at least 2xs/year) parent conferences, does it provide the family with a report of their child's developmental progress. Y/N __(5.3) (Score 1 point). Go to Number 4.
- 4) All these interactions are done in a culturally and linguistically appropriate way representing the parents being served. Y/N ____(5.4) (Score 1 point)

Scoring for PQI5:

Add up the total points based on the Ys; this will range from "0" to "4". The only way a program can receive a "4", is if a program has regularly scheduled parent conferences at least 2xs/year and provides the family with a report of their child's progress; and it is done in a culturally and linguistically appropriate way.

Record the number of points: _____(Range: 0

- 4) Total Score for Part 1 = _____

PART 2 - OBSERVATIONS:

INDICATOR 6): Educators Encourage Children to Communicate (Preschool Class)

Assessors will need to observe this item when they do their classroom observations. Initially you can ask educators or the director how children are encouraged to communicate but in order to gather reliable and valid information regarding this question/standard, it needs to be observed in the various interactions between staff and children. Things to look for would be more back and forth conversations rather than one-way conversations where educators are telling children what to do. Look for opportunities where children can describe what they are doing, how they feel about what they are doing, and why they are doing particular activities. Educators expand upon children's conversation.

These opportunities can occur anywhere in the classroom or outside, such as in dramatic play, tabletop activities or on the playground. Materials should be present that encourage communication such as toy telephones, puppets, flannel boards, dolls and dramatic play props, small barns, fire stations, or dollhouses. These create a lot of conversation among children as they assume many different roles. Children also talk when there is an interested person who listens to them. The staff in a high-quality early childhood classroom will use both activities and materials to encourage growth in communication skills.

How to Measure:

Observe the classroom for a minimum of 15 minutes. Once completed, consider where the classroom falls based on the following scale;
Score the classroom a 1 if the following occur:

- No activities used by staff with children to encourage them to communicate, for example: non talking about drawings, dictating stories, sharing ideas at circle time, finger plays, singing songs. Y/N _____(6.1)

- Very few materials accessible that encourage children to communicate. Y/N _____(6.2)

Score the classroom a 2 if the following occur (If the classroom does not have all 3 indicators but has 2 of the indicators then score this item 1+):

- Some activities are used by staff w/children to encourage them to communicate. Y/N _____(6.3)
- Some materials are accessible to encourage children to communicate. Y/N _____(6.4)
- Communication activities are generally appropriate for the children in the group. Y/N _____(6.5)

Score the classroom a 3 if the following occur (If the classroom does not have both indicators but has one of the indicators then score this item 2+):

- Communication activities take place during both free play and group times, for example: child dictates story about painting; small group discusses trip to store. Y/N __(6.6)
- Materials that encourage children to communicate are accessible in a variety of interest centers, for example: small figures and animals in block area; puppets and flannel board pieces in book area; toys for dramatic play outdoors or indoors. Y/N ____(6.7)

Score the classroom a 4 if the following occur (If the classroom does not have both indicators but has one of the indicators then score this item 3+):

- Staff balance listening and talking appropriately for age and abilities of children during communication activities, for example: leave time for children to respond; verbalize for child with limited communication skills. Y/N __(6.9)
- Staff link children's spoken communication with written language, for example: write down what children dictate & read it back to them; help them write notes to parents. Y/N _____(6.10)

Scoring for PQI 6:

Total up the number of "Y's" and record the appropriate level. In order for a classroom to receive a particular score, all "Y's" must be checked for the appropriate level (1 - 4) from above or partial credit given in order to obtain a "+". If there is a "+" please also mark it in the box.

Circle the Appropriate Level	***1***	***2***	***3***	***4***

INDICATOR 7): Infant Toddler Observation (if applicable) **(Infant Classroom)**

Conversations and questions should be used with all children, even young infants. Conversations using verbal and nonverbal turn-taking should be considered when scoring. Most conversations and questions initiated by infants will be nonverbal, such as widening of baby's eyes or waving arms and legs. Observe staff response to such nonverbal communication. For infants and toddlers, the responsibility for starting most conversations and asking questions belongs to the staff. As children become more able to initiate communication, staff should modify their approach in order to allow children to take on a greater role in initiating conversations and asking questions. Staff should provide answers to questions used by children if children cannot answer, and as children become more able to respond, questions should start to include those that the child can answer. If there was not an infant classroom, skip this Indicator and please note that here and on the summary score sheet by marking N/A: __

How to Measure:

Observe the classroom for a minimum of 15 minutes. Once completed, consider where the classroom falls based on the following scale;

Score the classroom a 1 if the following occurs:

- Staff never initiate turn-taking conversations with children, for example: rarely encourage baby to babble back; simple back and forth exchanges with verbal children never observed. Y/N _____(7.1)
- Staff questions are often not appropriate for children, or no questions are asked, for example: too difficult to answer; carry a negative message. Y/N _(7.2)
- Staff respond negatively when children can't answer questions, for example: "You should know this"; "You did not listen". Y/N _(7.3)

Score the classroom a 2 if the following occurs (If the classroom does not have all 3 indicators but has 2 of the indicators then score this item 1+):

- Staff sometimes initiate conversations with children, for example: babble back and forth with baby; copy baby's sounds; respond to baby's crying with verbal response; have short back and forth toddler interactions. Y/N _(7.4)
- Staff sometimes ask children appropriate questions and wait for the child to respond, for example: ask baby if she likes toy and pay attention as baby smiles; ask toddler what he is eating and wait for him to think of word. Y/N _(7.5)
- Staff respond neutrally or positively to children who can't answer questions. Questions asked are sometimes meaningful to children, for example: child responds with interest; does not ignore staff questions. Y/N ___(7.6)

Score the classroom a 3 if the following occurs (If the classroom does not have all 4 indicators but has 2 or more of the indicators then score this item 2+):

- Staff initiate engaging conversations with children throughout the observation, for example: show enthusiasm; use tone that attracts child's attention. Y/N _(7.7)
- Staff often personalize questions and/or conversations for individual children, for example: talk about children's families, preferences, interests; what they are playing with; what they did over weekend; child's mood; use child's name. Y/N __(7.8)

- Staff often pay attention to children's questions, verbal or nonverbal, and answer in a satisfying manner for the child. Y/N ____(7.9)
- Staff ask questions in which children show interest in answering, for example: make the questions funny or mysterious; use attractive tone; meaningful and not too difficult to answer. Y/N ___(7.10)

Score the classroom a 4 if the following occurs (If the classroom does not have both indicators but has one of the indicators then score this item 3+):

- Staff frequently have turn taking conversations with children throughout the observations. Many appropriate questions are used throughout the observation, during both play and routines. Y/N _____(7.11)
- Staff ask children appropriate questions, wait a reasonable time for child response, and then answer if needed, for example: "Are you hungry? . . . Yes, you are!"; "Where's the ball? . . . These it is! You found the ball". Y/N ___(7.12)

Scoring for PQI 7:

Total up the number of "Y's" and record the appropriate level. For a classroom to receive a particular score, all "Y's" must be checked for the appropriate level (1 - 4) from above or partial credit given in order to obtain a "+".

Circle the Appropriate Level	***1***	***2***	***3***	***4***

INDICATOR 8): Educators Use Language to Develop Reasoning Skills (Preschool)

Assessors will need to observe very carefully as this standard can be difficult to determine because it is tying language and cognition together. Again, this opportunity can occur in any setting in or out of the classroom because it is the basis for problem solving through the use of language. Also look for educators redirecting children's conversations when appropriate. Staff should use language to talk about logical relationships using materials that stimulate reasoning. Through the use of materials, staff can demonstrate concepts such as same/different, classifying, sequencing, one-to-one correspondence, spatial relationships, and cause and effect.

How to Measure:

Observe the classroom for a minimum of 15 minutes. Once completed, consider where the classroom falls based on the following scale;

Score the classroom a 1 if the following occur:

- Staff do not talk with children about logical relationships, for example: ignore children's questions and curiosity about why things happen, do not call attention to sequence of daily events, differences and similarity in number, size, shape, cause and effect. Y/N __ (8.1)
- Concepts are introduced inappropriately, for example: concepts too difficult for age and abilities of children, inappropriate teaching methods used such as worksheets without any concrete experiences; teacher gives answers w/o helping children to figure things out. Y/N _____ (8.2)

Score the classroom a 2 if the following occur (If the classroom does not have both indicators but has one of the indicators then score this item 1+):

- Staff sometimes talk about logical relationships or concepts, e.g.: explain that outside time comes after snacks, point out differences in sizes of blocks children use. Y/N __ (8.3)

- Some concepts are introduced appropriately for ages and abilities of children in group, using words and experiences, for example: guide children with questions and words to sort big and little blocks or to figure out why ice melts. Y/N _____(8.4)

Score the classroom a 3 if the following occur (If the classroom does not have both indicators but has one of the indicators then score this item 2+):

- Staff talk about logical relationships while children play with materials that stimulate reasoning, for example: sequence cards, same/different games, size and shape toys, sorting games, numbers and math games. Y/N _(8.5)
- Children are encouraged to talk through or explain their reasoning when solving problems, for example: why they sorted objects into different groups, in what way two pictures are the same or different. Y/N _____(8.6)

Score the classroom a 4 if the following occur (If the classroom does not have both indicators but has one of the indicators then score this item 3+):

- Staff encourage children to reason throughout the day, using actual events and experiences as a basis for concept development, e.g.: children learn sequence by talking about their experiences in the daily routine or recalling the sequence of a cooking project. Y/N _____ (8.7)
- Concepts are introduced based upon children's interests or needs to solve problems, for example: talk children through balancing a tall block building, help children figure out how many spoons are needed to set a table. Y/N _(8.8)

Scoring for PQI 8:

Total up the number of "Y's" and record the appropriate level. In order for a classroom to receive a particular score, all "Y's" must be checked for the appropriate level (1 - 4) from above or partial credit given in order to obtain a "+".

Circle the Appropriate Level	***1***	***2***	***3***	***4***

For quality key indicators 9 and 10 it is recommended that these be assessed/observed throughout the observation period and not just during key activity times. These two quality key indicators should be observed in two-minute blocks over ten sequences for a total of 20 minutes. These two items should also be used with each age group being assessed.

INDICATOR 9): Educators Listen Attentively When Children Speak

This quality indicator focuses on the early childhood educator(s) looking directly at the children with nods, rephrasing their comments, and engaging in conversations. Children should have the undivided attention of the specific educator they are addressing. Educators should not be looking away or pre- occupied with others. They should be at the child's level making eye contact. The intent is to observe all children and educators in the room.

How to Measure:

Do this in timed 2-minute observations recording each time you observe this occurring. Record at least 10 different observation periods. These do not need to be consecutive in order to fully observe classrooms and educators. Please use the following scale to assess your recordings: Likert Scale (1-4) where 1 = Never/Not at All; 2 = Somewhat/Few Instances; 3 = Quite a Bit/Many Instances; 4 = Very Much/Consistently):
Make the actual recordings using the Likert Scale (1-4) above for each individual observation and record in each cell below.

09 Observations:

09.1	2	3	4	5	6	7	8	9	09.10

Scoring for PQI 9:

Once all the observations are made, add up the results from the Likert Scale (1-4) and record the total number here: _____(Range: 10 - 40)(Divide this result by 10) = _____(1-4)(Round upward or downward to the whole number (3.7 = 4; 2.2 = 2)).

Circle the Appropriate Level	***1***	***2***	***3***	***4***

INDICATOR 10): Educators Speak Warmly to Children

This quality indicator focuses on the early childhood educator(s) always engaging in a caring voice and body language with every child. Educators do not use harsh language or commands in speaking to
children, but rather again are on the child's level making eye contact. Think of the way Fred Rogers would engage his audience where you always felt you were the most important person in the world when he talked to the TV.

How to Measure:

Do this in timed 2-minute observations recording each time you observe this occurring. Record at least 10 different observation periods. Please use the following scale to make your recordings: (This item is on a Likert Scale (1-4) where 1 = Never/Not at All; 2 = Somewhat/Few Instances; 3 = Quite a Bit/Many Instances; 4 = Very Much/Consistently):
Make the actual recordings using the Likert Scale (1-4) above for each individual observation and record in each cell below.
10 Observations:

10.1	2	3	4	5	6	7	8	9	10.10

Scoring for PQI 10:

Once all the observations are made, add up the results from the Likert Scale (1-4) and record the total number here: _____(Range: 10 - 40) (Divide this result by 10) = _____(1-4). (Round upward or downward to the whole number (3.7 = 4; 2.2 = 2)).

Circle the Appropriate Level	***1***	***2***	***3***	***4***

Results

This section provides the resultant scoring protocol that is generated from the above methods: Contact Hour Metric and the Program Quality Indicators. It is followed by the opening screen to the CCEEHM App.

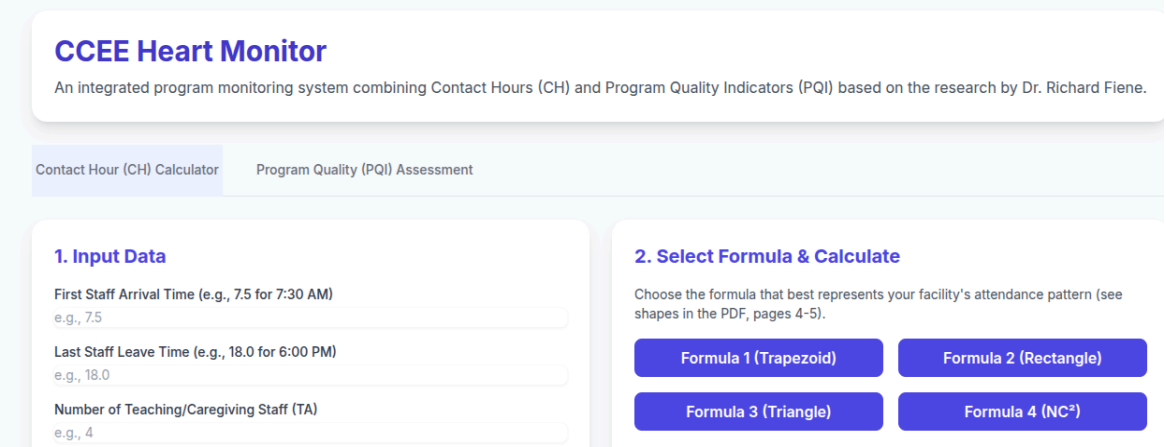
This scoring protocol (Table 2) which is generated from the Program Quality Indicators AI algorithms rank orders all programs on how well the classrooms measure up during the AI observations. The standardized scores are to the left while the actual classroom scores are to the right in the below table. The levels are specified as high quality, high-mid quality, mid-low quality, and low quality based upon the actual scores obtained.

Table 2: Program Quality Indicators Artificial Intelligence (PQIAI) Scoring Protocol

LEVEL	Standardized Scores	Actual Scores
High Quality	**Mixed Age: 36+** **Preschool: 32+** **Infant-Toddler: 28+**	Mixed Age: _____ Preschool: _____ Infant-Toddler: _____
High - Mid Quality	**Mixed Age: 30 – 35** **Preschool: 26 - 31** **Infant-Toddler: 22 - 27**	Mixed Age: _____ Preschool: _____ Infant-Toddler: _____
Mid – Low Quality	**Mixed Age: 20 – 29** **Preschool: 16 - 25** **Infant-Toddler: 12 - 21**	Mixed Age: _____ Preschool: _____ Infant-Toddler: _____
Low Quality	**Mixed Ages: 19 or less** **Preschool: 15 or less** **Infant-Toddler: 11 or less**	Mixed Age: _____ Preschool: _____ Infant-Toddler: _____

This is the opening screen to the Child Care and Early Education Heart Monitoring App (CCEEHM):

Figure 2



This is the CCEE Heart Monitor Application: The ***Child Care and Early Education Integrated Program Monitoring System.*** It has two main sections, accessible through tabs:

1. Contact Hour (CH) Calculator: Input your facility's operational data to calculate the Contact Hour metric, which helps in analyzing structural quality. You can also include square footage for an expanded calculation.
2. Program Quality (PQI) Assessment: Go through the 10 indicators to evaluate the process quality of an early education program. The tool will automatically score each indicator and provide a final quality level based on the age group you select.

The Discussion section contains a side by side comparison of two hypothetical programs, one of high quality and one of low quality. This case study gives the details of what the results would look like in utilizing the CCEEHM App.

Discussion

Comparative Analysis of Program Quality: A High vs. Low-Quality Child-Care Program Assessment Using the CCEEHM Framework

1. Introduction: A Tale of Two Programs

This report presents a detailed, side-by-side comparison, typical of a comprehensive program review, of two hypothetical early childhood programs—one high-quality and one low-quality—using the Child Care and Early Education Heart Monitor (CCEEHM) framework. The purpose of this analysis is to illustrate how the CCEEHM's integrated approach provides a comprehensive and objective measure of a program's overall effectiveness and its environment for children. By combining metrics for both structural and process quality, the framework moves beyond simple compliance checklists to create a holistic picture of the daily experiences that shape early development.

The CCEEHM framework is built on two core components, which will be used to structure this comparative analysis:

- **Structural Quality:** This foundational element is measured by the **Contact Hour (CH)** metric. It focuses on health, safety, and regulatory compliance elements and serves as a more effective and efficient metric for measuring compliance with standards like adult-child ratios and group sizes.
- **Process Quality:** This component is measured by the **Program Quality Indicators (PQI)**. It assesses the "heart" of quality—the daily interactions, curriculum, and learning environment that directly impact a child's developmental experience and well-being.

This report will begin by analyzing the foundational element of structural quality, demonstrating how the Contact Hour metric distinguishes a safe, compliant program from an unsafe, non-compliant one.

2. Structural Quality Analysis: The Contact Hour (CH) Metric

The Contact Hour (CH) metric is a strategically important measure of a program's foundational quality. It is not merely an administrative number but a direct assessment of regulatory compliance with adult-child ratios and group sizes. These standards are critical for ensuring child safety, providing adequate supervision, and creating an environment where meaningful interactions can occur. This section will demonstrate how the CH metric quantitatively differentiates a well-managed, compliant program from an overpopulated, non-compliant one.

To illustrate this, we will examine two hypothetical preschool classrooms: **Program A (High Quality)** and **Program B (Low Quality)**. Let's assume the required adult-child ratio for this age group is 1 educator for every 10 children (1:10).

CH Metric Question	Program A (High Quality) Response	Program B (Low Quality) Response
1. When does the facility open?	8:00 AM	8:00 AM
2. When does the facility close?	4:00 PM	4:00 PM
3. Number of teaching staff (TA)?	2	2
4. Number of children (NC)?	18	24
5. When does the last child arrive?	8:00 AM	8:00 AM
6. When does the first child leave?	4:00 PM	4:00 PM

For this analysis, we assume a scenario where full enrollment is present for the entire day to align with the reference model, using the formula CH = (NC x TO) / TA.

Contact Hour (CH) Calculation

- **Program A (High Quality)**
  - Number of Children (NC) = 18
  - Total Hours Open (TO) = 8
  - Total Teaching Staff (TA) = 2
  - **Calculation:** (18 * 8) / 2 = 72
  - **Final CH Value: 72**
- **Program B (Low Quality)**
  - Number of Children (NC) = 24
  - Total Hours Open (TO) = 8
  - Total Teaching Staff (TA) = 2
  - **Calculation:** (24 * 8) / 2 = 96
  - **Final CH Value: 96**

Analysis of Results

To determine compliance, we compare each program's calculated CH value against the maximum allowed CH value for their number of children at the required 1:10 ratio, as specified in the *Contact Hour (CH) Conversion Table*.

- **Program A (High Quality):** With 18 children and 2 staff, Program A maintains a 1:9 ratio, which is better than the required 1:10. According to the CH Conversion Table, for 18 children (NC) at a required 1:10 ratio, the maximum allowed CH value is **72**. Program A's calculated CH of **72** is equal to this compliance threshold.
    - **Conclusion:** Program A is **in compliance** with adult-child ratio and group size standards.
- **Program B (Low Quality):** With 24 children and only 2 staff, Program B operates at a 1:12 ratio, exceeding the required 1:10 limit. To determine the maximum allowed CH for a compliant program with 24 children at a 1:10 ratio, one must first calculate the required number of staff (24 children / 10 = 2.4, which requires 3 staff members). The compliant CH is therefore (24 children * 8 hours) / 3 staff = 64. This value of 64 represents the compliance threshold in the CH Conversion Table. Program B's calculated CH of **96** significantly exceeds this threshold.
    - **Conclusion:** Program B is **out of compliance**. The program is characterized as "overpopulated." Its density display would resemble the final example from the source text, a depiction that "clearly indicates a very high CH and non-compliance with ACR and GS."

This quantitative analysis reveals a critical distinction: Program A provides a safe and compliant structural foundation, whereas Program B does not. This structural failure creates an environment where high-quality interactions and individualized attention—the core components of process quality—are nearly impossible to achieve. We now turn to the qualitative assessment of the program environment to see how this foundation impacts the daily experiences of children.

3. Process Quality Analysis: The Program Quality Indicators (PQI)

The Program Quality Indicators (PQI) are the tools used to measure the "heart" of quality—the developmental and interactional experiences that define a child's day. While structural metrics ensure safety, the PQIs evaluate the richness of the curriculum, the warmth of interactions, and the overall supportiveness of the learning environment. This section will systematically compare Program A and Program B across all 10 PQIs to reveal the profound differences in their approaches to early care and education.

PQI 1: ECE III Educators (Credentialed Staff)

This indicator measures the percentage of teaching staff with higher-level credentials in early childhood education (ECE III), which is linked to higher quality interactions and curriculum implementation.

Program A (High Quality)	Program B (Low Quality)
A review of staff records shows that **8 out of 10** teaching staff (80%) are certified at the ECE III level. This high concentration of qualified educators suggests a strong commitment to professional knowledge and practice. **Score: 4** (76-100%)	A review of staff records shows that only **2 out of 10** teaching staff (20%) are certified at the ECE III level. This low percentage indicates a significant gap in staff qualifications and training. **Score: 1** (0-25%)

PQI 2: Stimulating and Dynamic Environment

This indicator uses an 11-point checklist to assess whether the classroom environment is child-centered, accessible, and reflective of children's interests and cultures.

Program A (High Quality)	Program B (Low Quality)
An observer marks 'Yes' for **10 of the 11** checklist items. Children freely access a variety of authentic materials, their work is displayed respectfully, and documentation of their learning projects is evident throughout the room. **Score: 4** (91%)	An observer marks 'Yes' for only **2 of the 11** checklist items. Materials are stored out of children's reach, the physical environment lacks evidence of child-authored work, and there are no family photos or sufficient print materials available. **Score: 1** (18%)

PQI 3: Developmentally Appropriate Curriculum

This indicator assesses whether the program uses individual child assessments to inform an emergent, developmentally appropriate curriculum.

Program A (High Quality)	Program B (Low Quality)
A review of 10 children's records reveals that **9 records (90%)** show a clear, documented link between developmental assessments and an individualized curriculum plan. All three key elements (emergent curriculum, co-learning,	A review of 10 children's records finds only **1 record (10%)** showing a link between assessment and curriculum. Most files lack developmental assessments or show a

and documented planning) are present. **Score: 4** (90%)	"canned," one-size-fits-all curriculum with no individualization. **Score: 1** (10%)

PQI 4: Opportunities for Staff and Families

This indicator measures the presence of policies and practices that foster strong, respectful, two-way communication and relationships between staff and families.

Program A (High Quality)	Program B (Low Quality)
The program has implemented all three examples of strong family communication. It provides materials in diverse formats, uses multiple modes of two-way communication (e.g., apps, regular calls), and demonstrates respectful engagement with family goals. **Score: 4** (100%)	The program has no formal policies for family communication. Communication is sporadic, one-way (e.g., occasional flyers), and does not actively engage families in a partnership. **Score: 1** (0%)

PQI 5: Information on Child's Progress

This indicator evaluates the formality and frequency with which programs share information about a child's developmental progress with their family.

Program A (High Quality)	Program B (Low Quality)
The program conducts formal parent-teacher conferences at least twice a year, provides detailed written progress reports, and ensures all communication is done in a culturally and linguistically appropriate manner for every family. **Score: 4**	The program offers neither regularly scheduled conferences nor written progress reports. Information is shared only if a problem arises. **Score: 0**

PQI 6: Educators Encourage Children to Communicate (Preschool)

This observational indicator assesses how effectively educators use materials and interactions to foster children's communication skills.

Program A (High Quality)	Program B (Low Quality)
During observation, educators skillfully balance listening and talking, leaving ample time for children to respond. They link children's spoken ideas to written language by	Observers note that no specific activities or materials are used to encourage communication. Staff-child talk consists primarily of one-way commands, and

writing down their stories and reading them back. **Score: 4**	children's attempts to converse are often ignored. **Score: 1**

PQI 7: Infant Toddler Observation

This indicator is for infant/toddler rooms. For this preschool-focused comparison, it is marked as N/A.

PQI 8: Educators Use Language to Develop Reasoning Skills (Preschool)

This observational indicator measures whether educators use daily conversations and experiences to build children's logical thinking and problem-solving abilities.

Program A (High Quality)	Program B (Low Quality)
Educators consistently use daily events to develop concepts. For example, they help children sequence steps in a cooking project and encourage them to reason through how to build a stable block tower. **Score: 4**	Educators do not talk with children about logical relationships like cause-and-effect or sequencing. Concepts are introduced via worksheets without concrete experiences, and staff simply provide answers rather than helping children think. **Score: 1**

PQI 9: Educators Listen Attentively

This indicator uses a series of timed observations to measure how consistently educators give children their undivided attention when they speak.

Program A (High Quality)	Program B (Low Quality)
Across 10 timed observations, educators consistently make eye contact, nod, and rephrase children's comments. Their average Likert score is **3.7**, demonstrating a consistent pattern of attentive listening. **Score: 4** (rounded up from 3.7)	Across 10 timed observations, educators rarely listen attentively, often looking away or continuing other tasks while a child is speaking. Their average Likert score is **1.3**. **Score: 1** (rounded down from 1.3)

PQI 10: Educators Speak Warmly

This indicator uses timed observations to measure the emotional tone of educator-child interactions, focusing on caring voice and body language.

Program A (High Quality)	Program B (Low Quality)
Observations reveal that educators consistently speak to children with a caring, respectful tone and warm body language. Their average Likert score is **3.8**, indicating a consistently positive emotional climate. **Score: 4** (rounded up from 3.8)	Observations show that educators rarely speak warmly. Their tone is often flat, harsh, or dismissive. Their average Likert score is **1.4**. **Score: 1** (rounded down from 1.4)

The detailed analysis of each indicator reveals a consistent pattern of high performance in Program A and deficient performance in Program B. The final section will synthesize these scores into an overall quality classification.

4. Final Scoring and Quality Classification

This section synthesizes the individual Program Quality Indicator (PQI) scores to generate a final, data-driven quality classification for each program. This culminating step demonstrates the CCEEHM's ability to provide a clear summary of overall process quality, transforming detailed observations into an actionable and easily understood rating.

The table below summarizes the scores for Program A and Program B across all applicable indicators.

Program Quality Indicator (PQI)	Program A: High Quality Score	Program B: Low Quality Score
PQI 1: ECE III Educators	4	1
PQI 2: Stimulating and Dynamic Environment	4	1
PQI 3: Developmentally Appropriate Curriculum	4	1
PQI 4: Opportunities for Staff and Families	4	1
PQI 5: Information on Child's Progress	4	0
PQI 6: Educators Encourage Communication	4	1

PQI 7: Infant Toddler Observation	N/A	N/A
PQI 8: Educators Use Language for Reasoning	4	1
PQI 9: Educators Listen Attentively	4	1
PQI 10: Educators Speak Warmly	4	1
Total Score (Preschool)	**36**	**8**

Analysis and Classification

Using the *Program Quality Indicators Artificial Intelligence (PQIAI) Scoring Protocol*, we can assign a final classification to each program based on its total score.

- **Program A (High Quality):** With a total score of **36**, Program A significantly exceeds the preschool threshold of 32 or higher.
    - **Classification: High Quality**
- **Program B (Low Quality):** With a total score of **8**, Program B falls well below the preschool threshold of 15 or less.
    - **Classification: Low Quality**

These classifications provide a definitive summary of the vast differences in process quality between the two programs, which will be discussed in the report's conclusion.

5. Conclusion: The Integrated Picture of Quality

The comparative analysis of Program A and Program B using the CCEEHM Scoring Protocol framework reveals a stark contrast that extends across every dimension of quality. The synthesis of both structural (Contact Hour) and process (Program Quality Indicators) metrics paints a complete and compelling picture of two vastly different environments for young children.

The key differentiators are clear. **Program A** not only demonstrates regulatory compliance and safety with a proper CH score but also excels in creating a rich, supportive, and developmentally appropriate environment, as evidenced by its high PQI score. It is a program where a safe foundation enables high-quality interactions, intentional teaching, and strong family partnerships to flourish.

In contrast, **Program B** is failing on all fronts. It is structurally non-compliant and unsafe, operating with an overpopulated classroom reflected in its high CH score. This foundational failure is mirrored in its process quality, where a low PQI score indicates an environment lacking

qualified staff, a meaningful curriculum, and the warm, responsive interactions that are essential for positive child development.

Ultimately, this analysis reinforces the value of the CCEEHM framework as an integrated system that moves the field beyond isolated compliance checks. While a simple licensing visit might check a program's pulse, the CCEEHM provides an EKG of its heart—measuring not just the structural factors that keep children safe, but the vital process quality interactions that make their hearts and minds grow.

Conclusion

The CCEEHM is an example of an integrated program monitoring system that puts structural and process quality on the same platform, something that has not been done in the early care and education field. This paper has delineated how to do this by starting with the innovative Contact Hour (CH) Metric and then combining that methodology with the Program Quality Indicators (PQI) generated utilizing the Key Indicator Methodology (KIM)(Fiene & Nixon, 1985) in which each CH is given a process quality score as described in the Results section. The PQIs are drawn from early care and education accreditation systems, professional development systems, and quality rating & improvement systems.

Integrated program monitoring systems build from differential monitoring systems approach which utilizes risk assessment and key indicator methods as their focal point. WIth integrated program monitoring systems program quality is infused into the rule making which enhances the structural level of quality by having a process quality element building upon the structural quality foundation. The CCEEHM has all these key elements built into its data analytical architecture based upon artificial intelligence and big data analysis.

Another way of looking at this framework is through the lens of cognitive computing in which the measurement strategy goes from nominal measurement which is predominant with structural quality to more of an ordinal measurement strategy with the introduction of process

quality elements. This leads to program quality indicators being measured at the ordinal level and ultimately leads to a Regulatory Compliance Scale (RCS) which has been proposed as an innovation for regulatory science measurement (Fiene, 2025b).

The CCEEHM could not be built without AI and big data analysis given the tremendous number of observations that need to be made in order to build and standardize the Scoring Protocol (Table 2) in the Results section. The number of observers it would take in order to build such a Scoring Protocol would have been prohibitively expensive. Also, the professional development and mentoring/coaching linkages to these AI Observations are extensive and can be used in programs to make improvements in their teaching staff.

Integrated program monitoring is a cost effective and efficient approach which protects children from harm while at the same time enhancing their daily experiences in the classroom. It attempts to provide the best of both worlds in "doing no harm" while "doing good", the twin pillars of an early care and education regulatory and quality framework (Fiene, 2025a). It also has the ability to inform professional development and technical assistance systems by building effective and efficient coaching/mentoring interventions based upon the results of the CCEEHM Scoring Protocol. This is a major enhancement in being able to link the program monitoring system to the professional development quality initiative.

The CCEEHM App is an innovation that should go a long way in improving the structural and process quality elements of early care and education programs in a cost effective and efficient manner. The Integrated Program Monitoring Systems Approach is the latest iteration in a long development of program monitoring systems from Uniform Program Monitoring to Instrument Based Program Monitoring to Differential Program Monitoring.

The computer generated hypothetical report highlighted in the Discussion section clearly demonstrates how the CCEEHM App will produce results that clearly differentiate between high

and low levels of quality in early care and education programs. It will be interesting to see how this new technology is used throughout the early care and education field in assessing both structural and process quality in children's classrooms.

Funding

This research received no external funding.

Conflicts of Interest

The author declares no conflict of interest.

References

Fiene, R. (2025a). Potential Solution to the Child Care Trilemma Revisited: Finding the "Right Rules" - The Holy Grail of Early Care and Education, *Exchange*, Summer 2025, pps 64-66.

Fiene, R. (2025b). Finding the Rules that Work: An emerging paradigm promised to close the gap between regulatory compliance scores and quality of childcare services. *American Scientist, 113, 16-21, (January-February 2025).*

Fiene, R. (2025c). The Uncertainty–Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies. *Knowledge*, *5*(2), 8. https://doi.org/10.3390/knowledge5020008

Fiene, R. (2024). *Regulatory Compliance and Monitoring Systems*. Research Institute for Key Indicators and National Association for Regulatory Administration.

Freer, R. & Fiene, R. (2023). Regulatory Compliance and Quality Programs: Constraints and Opportunities for Integration. *Journal of Regulatory Science*, *11*(1), 1-10. https://doi.org/10.21423/JRS.REGSCI.111264

Fiene, R. (2022). Regulatory compliance monitoring paradigms and the relationship of regulatory compliance/licensing with program quality: A policy commentary. *Journal of Regulatory Science 10:1–7*.

Fiene, R. & Stevens, S. (2021). *Contact Hours as a New Metric Replacing Group Size and Staff-Child Ratios as well as a New Metric for COVID19 Thresholds.* Research Institute for Key Indicators/Penn State.

Fiene, R. (2019). A treatise on the theory of regulatory compliance*, Journal of Regulatory Science*, *Volume 7(1), pps 1-3.* (doi.org/10.21423/jrs-v07fiene).

Kontos, S. & Fiene, R. (1987). Child care quality, compliance with regulations, and children's development: The Pennsylvania Study, in *Quality in Child Care: What Does Research Tell Us?*, Phillips, editor, Washington, D.C.: National Association for the Education of Young Children.

Fiene, R. & Nixon, M. (1985). Instrument based program monitoring and the indicator checklist for child care, *Child Care Quarterly, 14*(3), 198-214.

Richard Fiene PhD, Senior Research Psychologist & Regulatory Scientist, Edna Bennett Pierce Prevention Research Center, Pennsylvania State University, https://prevention.psu.edu/person/rick-fiene/

nara National Association for Regulatory Administration

KEY QUALITY INDICATORS

A white paper to introduce the concepts and application of key quality indicators in human care regulatory environments

43 Town & Country Drive
Suite 119, #121
Fredericksburg, VA 22405
888-674-7052
www.naralicensing.org
U.S. Federal ID No. 54-1542015

CONSUMER PROTECTION THROUGH PREVENTION

Contents

Overview 2
Introduction to the Theory and Methodology 2
The Original Model: 3
Theory of Regulatory Compliance: 4
Differential Monitoring and the Key Indicator Methodology: 4
The Emergence of the Regulatory Compliance Scale: 6
The Various Quality Systems: Accreditation, Professional Development, Training, & Technical Assistance, Quality Observations, and Quality Rating & Improvement Systems (QRIS), and the Development of a Quality Indicators Scale 6
Conclusion and Expansion Beyond Child Care 7
About the Author: 8
Contributors 10
References and Future Readings: 11

NARA White Paper on Quality Indicators

Overview

This paper introduces the concept and application of Quality Indicators in the Child Care and Early Education (CCEE) field. It begins by delineating how quality has historically been categorized into two dimensions: structural quality and process quality, as established in the research literature (Morgan, 1979; Peisner-Feinberg et al., 2001; Duncan, 2003). The paper then examines a range of models and systems to assess their effectiveness in achieving high-quality child care.

A central focus is the Theory of Regulatory Compliance, which offers a unifying framework for linking quality and regulatory compliance. This theory supports the development of a key indicator methodology that has been applied extensively across the United States and Canada (Fiene, 2019).

Building on this conceptual foundation, the paper highlights several systems that contribute to child care quality, including accreditation, professional development, training, technical assistance, Quality Rating and Improvement Systems (QRIS), and observational assessments such as the Environmental Rating Scales (ERS) (Fiene, 2022). Each of these systems is examined for its contribution to identifying key quality indicators.

Following this analysis, the paper introduces a consolidated Quality Indicators Tool that has evolved from these multiple systems (Fiene, 2024). The concluding section considers how this integrated approach to quality and compliance can be applied to other human service fields beyond early childhood care and education.

Introduction to the Theory and Methodology

Child care and early education (CCEE) quality has been defined in the research literature along a continuum between structural and process quality. Structural quality refers to countable, objective standards, while process quality reflects the more nuanced interpersonal dynamics between adults and children.

Structural quality is typically associated with licensing rules that safeguard children's health and safety in out-of-home care. Process quality, on the other hand, captures the nature of teacher-child interactions. These are often assessed through quality observation tools such as the Environmental Rating Scales (ERS), commonly used within Quality Rating and Improvement Systems (QRIS) (Peisner-Feinberg et al., 2001; Duncan, 2003).

Process quality is often considered the core of early childhood quality. It captures the intricate details of what happens in classrooms during individual and group interactions. This includes positive and engaging language, emotional climate, and opportunities for children to solve problems. Process quality emphasizes how well teachers facilitate these experiences, either through direct engagement or by arranging the learning environment to support them.

Structural quality refers to surrogate measures such as compliance with staff-child ratios, group sizes, or the number of regulatory violations. It does not typically assess emotional tone, instructional style, or classroom atmosphere. Occasionally, it may include aspects of curriculum, but its primary focus is on health and safety standards; those factors intended to prevent harm rather than promote enrichment. Enhancing the child's learning environment is primarily addressed through process quality.

From an organizational perspective, structural quality is most commonly found in licensing rules and regulations. Process quality, by contrast, is captured through tools such as the Environmental Rating Scales (ERS) or the Classroom Assessment Scoring System (CLASS). Although these tools were originally designed for standalone use, they are now more commonly embedded within broader quality initiatives. ERS is often used within QRIS, and CLASS is widely used in Head Start.

Structural and process quality are best understood as complementary. Structural quality provides a foundational layer, while process quality builds upon that base to support child development more holistically.

Another useful way to conceptualize quality is through the image of a spectrum. Structural and process quality can be placed at opposite ends of a continuum that reflects the range of quality interventions developed over the past 40 years. Interventions that align with structural quality include licensing, QRIS, Head Start Performance Standards, accreditation, and professional development systems. Process quality is primarily represented by tools such as ERS and CLASS.

This spectrum can be imagined as a prism: instead of separating light into colors, it separates the dimensions of quality into distinct intervention systems, each contributing in a different way to the overall goal of high-quality care.

The Original Model:

Gwen Morgan introduced the concept of a quality spectrum in a 1979 article published in Young Children (Morgan, 1979), which examined the components of child care quality. In that article, quality was categorized into two broad system types: regulatory and non-regulatory.

Regulatory components included licensing, contracting, best practices, credentialing, rate setting, and accreditation. Non-regulatory components included professional development and training systems, referral and resource agencies, advocacy organizations, and public education.

The Morgan Model was one of the earliest efforts to integrate these systems into a unified framework for understanding child care quality. Since its introduction, the model has expanded to include additional systems that have emerged over time, such as new accreditation frameworks, Quality Rating and Improvement Systems (QRIS), and statewide professional development and technical assistance initiatives—particularly those that use coaching or mentoring approaches.

The next section introduces a licensing and regulatory science framework that depicts the relationship between regulatory compliance and program quality.

Theory of Regulatory Compliance:

The Theory of Regulatory Compliance (TRC) (Fiene, 2019) marked a major breakthrough in how key indicators are conceptualized and applied in the licensing and monitoring of child care quality. It has been described as a paradigm shift that moved the licensing field and regulatory science away from a uniform monitoring approach toward a differential monitoring model (Fiene, 2025c,d). This transition enabled the development of the key indicator methodology, which made abbreviated inspections possible. Without TRC, the traditional model of uniform program monitoring might have remained the standard for conducting all child care inspections - but in practice, it proved insufficient.

Uniform monitoring was the dominant approach for licensing and oversight through the 1970s and 1980s. However, as evidence accumulated, it became clear that a more targeted method was needed. This shift laid the groundwork for the application of key indicators in a more nuanced and predictive way.

TRC serves as an overarching framework that explains how structural, and process quality interacts. One of its most important insights is the value of substantial compliance with structural quality rules. This was demonstrated through the identification of a ceiling effect when comparing structural and process quality across various systems, including licensing, Head Start, accreditation, and QRIS. Among these, licensing exhibited the most pronounced ceiling effect and, in some cases, a diminishing returns pattern when moving from substantial compliance to full 100 percent compliance. Although all structural systems showed similar tendencies, process quality followed a different pattern, typically a linear relationship with a normal data distribution. In contrast, structural quality data were non-linear and positively skewed. These contrasting patterns have been confirmed repeatedly in CCEE research over the past five decades.

As a result, TRC has helped reframe how licensing decisions are made. It recognizes that substantial compliance may be sufficient for issuing a full license, and in some cases, may be more predictive of program quality than full compliance. This perspective has also led to the use of abbreviated or targeted inspections that focus on key predictor rules and high-risk standards. This shift established the foundation for differential monitoring.

Importantly, differential monitoring allows quality indicators to be integrated into licensing and program oversight, something that was not easily achieved under uniform monitoring models. The structure and methodology of differential monitoring are described in the following section.

Differential Monitoring and the Key Indicator Methodology:

Recent policy studies (Freer & Fiene, 2023) confirm that quality indicators can be systematically identified and integrated into the licensing and regulatory framework. These indicators form the foundation of the differential monitoring model, serving as anchors for both structural and process quality.

Key indicators have repeatedly been shown to statistically predict overall compliance with the full set of rules, regulations, and standards through both ongoing research and analysis of key indicator system

effectiveness in states and provinces that have incorporated key indicator systems into their licensing programs. This predictive relationship has been repeatedly validated across systems such as licensing, QRIS, Head Start, accreditation, and ERS. It also supports the development of a new quality indicator scale that is based on the same statistical methodology but is applied to nonregulatory standards that measure quality in early childhood development and education using standardized quality measurement tools.

Originally developed for child care licensing (Fiene & Nixon, 1985), the key indicator methodology has since been successfully applied to the identification of key compliance indicators, key risk indicators, key performance indicators, and key quality indicators. It is essential that any use of this methodology follows the original framework outlined in Fiene's research and NARA's implementation guidelines.

Two methodological components are particularly critical: the weighting of rules and the dichotomization of data. These techniques represent intentional departures from conventional predictive models in the testing and measurement literature. When properly applied, they help mitigate or eliminate false positives and false negatives in licensing decisions, particularly when determining whether a facility should receive a license (Fiene, 2024).

From a statistical perspective, correlations between structural and process quality have occasionally been significant, but they are generally modest. This is largely due to differences in how the two types of data are structured and measured. Structural quality data tend to show a ceiling effect and a positively skewed distribution, whereas process quality typically follows a more linear relationship and exhibits a normal distribution. These distinct statistical profiles help explain why correlations between structural and process quality are often weak, even though each provides meaningful insight on its own.

A key factor underlying this divergence is the binary nature of structural quality measurement. Licensing rules are inherently dichotomous: a provider either complies with a given rule or does not. This yes/no structure limits the ability of structural data to capture variation in performance above the minimum threshold. Once compliance is achieved, there is no higher score available under the traditional model.

In contrast, process quality assessments use ordinal or interval scales to capture a broader continuum of performance from very low quality to exemplary practice. These tools can account for gradations in quality, such as partially met expectations, developing skills, or consistently high engagement. As a result, process quality tools are better equipped to differentiate among programs performing above the basic compliance level.

This difference in measurement design not only contributes to the ceiling effect in structural quality but also restricts its ability to meaningfully correlate with process quality across the full range of provider performance. Structural measures tend to group most providers at or near the top, making it difficult to distinguish between high and moderate performers. By contrast, process quality ratings offer a more nuanced and scalable view of practice.

Despite these limitations, both structural and process quality are generally effective at identifying the lowest-performing programs. Structural violations and poor-quality interactions tend to co-occur in these settings, reinforcing the value of using both types of data to flag high-risk environments

The Emergence of the Regulatory Compliance Scale:

Challenges in differentiating high-performing and average-performing programs prompted the development of a new structural quality metric: the Regulatory Compliance Scale (RCS) (Fiene, 2025b). This scale was introduced for two primary reasons. First, the RCS aligns with the Theory of Regulatory Compliance by emphasizing substantial compliance and incorporating a categorical scoring structure. Second, its ordinal format corresponds well with widely used process quality tools, which typically apply a 1–7 rating scale.

By establishing a categorical structure, the RCS enables structural quality to be analyzed on more equal footing with process quality from a statistical measurement standpoint. Pilot testing in several jurisdictions suggests that the RCS is a more effective comparative tool than traditional methods that rely solely on violation counts or frequency data drawn from rule, regulation, or standard compliance (Fiene, 2024).

While these ideas have been addressed individually in prior literature (Trivedi, 2015), this paper brings them together to demonstrate the broader impact of the Theory of Regulatory Compliance on both structural and process quality assessment. Future replication of these findings by research psychologists and regulatory scientists would support an important policy advancement: recognizing substantial compliance as a sufficient threshold for full licensure and embedding differential monitoring as a standard regulatory practice across the CCEE field.

To support such a shift, the ceiling effect associated with structural quality must be reliably replicated when compared to the more evenly distributed nature of process quality scores (Fiene, 2025d).

The Various Quality Systems: Accreditation, Professional Development, Training, & Technical Assistance, Quality Observations, and Quality Rating & Improvement Systems (QRIS), and the Development of a Quality Indicators Scale

The same methodology used to identify Key Compliance Indicators has also been applied to the identification of Key Quality Indicators, which are the focus of this section. This Key Indicator Methodology is adaptable to any system governed by rules, regulations, or standards. Its effectiveness has been demonstrated in child welfare (Fiene, 1987), foster care reviews (Stevens, Fiene, Blevins, & Salzer, 2020), and in the identification of Key Performance and Key Risk Indicators for the Head Start Monitoring System.

The first quality initiative examined is accreditation, which has evolved significantly in the child care and early education field. Some accreditation systems are based on expert consensus or literature review, while others rely on empirically validated research methods. This section focuses specifically on one

accreditation system developed using the key indicator approach discussed earlier (Fiene, 1995, 1996). The National Early Childhood Program Accreditation (NECPA) system was developed and field-tested in the early 1990s as a cost-effective, efficient alternative to existing models. Its standards overlapped with both the National Association for the Education of Young Children's (NAEYC) accreditation framework and the newly developed Caring for Our Children standards (AAP & APHA, 1992). In validation studies (Fiene, 1996), NECPA was found to be highly correlated with the more prominent NAEYC system.

The second major initiative involves professional development, training, and technical assistance systems, which have been implemented across all states using quality incentive funds from the Child Care Development Block Grant. Application of the key indicator methodology found that coaching and mentoring were more effective than traditional methods in improving teacher-child interactions (Fiene, 2002). As a result, many states have shifted from workshop-based training to a blended model emphasizing coaching and mentoring.

The third major quality initiative is the implementation of Quality Rating and Improvement Systems (QRIS) as a supplement to traditional licensing systems. QRIS has proven effective in improving both structural and, in many cases, process quality across participating states (Zellman & Fiene, 2012). Approximately half of all states have adopted QRIS as of this writing. Within QRIS frameworks, Key Quality Indicators (KQIs) have been developed, particularly in the domains of family engagement and communication (Fiene, 2014). These indicators are central to program scoring within QRIS systems. Programs that meet KQIs are statistically more likely to receive higher QRIS ratings, typically at Level 3 or 4.

The final initiative discussed is quality observation, typically conducted using Environmental Rating Scales (ERS) (Harms, Clifford, & Cryer, 2012). ERS tools are widely used in QRIS systems as well as in other quality measurement contexts. A research study was conducted to determine whether ERS data could yield identifiable Key Quality Indicators. The study found that subscales related to language exchange and reasoning skills between teachers and children served as strong predictors of overall ERS scores.

This leads to the current state of Key Quality Indicator development. Most recently, a new Key Quality Indicators tool and software application has been proposed (Fiene, 2025a). The tool integrates findings from the initiatives described above, combining structural and process quality metrics into a single platform intended to improve both efficiency and comprehensiveness. However, the tool is currently in beta testing and requires further empirical validation to determine its long-term utility in the CCEE field. If validated, it may reshape how structural and process quality are assessed: consolidating currently separate systems into a unified approach.

Conclusion and Expansion Beyond Child Care

Quality indicators have expanded the scope of program monitoring, moving the field from uniform monitoring to differential monitoring, and now toward integrated monitoring approaches that explicitly incorporate quality (Freer & Fiene, 2023). Historically, key indicator methodologies focused primarily on regulatory compliance, with quality dimensions often excluded from formal review processes. This has

changed with the identification of Key Quality Indicators (KQIs) drawn from accreditation systems, Quality Rating and Improvement Systems (QRIS), professional development and technical assistance initiatives, and observational tools such as the Environmental Rating Scales (ERS).

The resulting integrated monitoring model offers jurisdictions a promising framework for evaluating both structural and process quality in their inspection systems. It represents an opportunity to align compliance monitoring with broader goals related to service effectiveness and developmental outcomes, particularly in early childhood settings.

This paper aims to provide guidance on the research supporting KQIs and how these indicators can be used to predict overall quality across structural and process dimensions. While the examples are drawn from the Child Care and Early Education (CCEE) field, the underlying approach and methodology are applicable across a wide range of human service systems, including child residential and adult residential programs. Indeed, this methodology can be applied in any setting governed by rules, regulations, or standards as discussed earlier.

However, there are limitations to expanding the model into other human services contexts, particularly adult care. One significant challenge is the absence of established population-wide quality evaluation systems for older youth and adult service recipients. While evaluation methods do exist in these domains, they are often designed around the experience of the individual, assessing the quality of care received by a specific person rather than the overall performance of the service provider. This distinction complicates the development of system-wide Key Quality Indicators.

Additionally, it is inherently easier to define and measure growth and development in early childhood due to well-established developmental and educational milestones. In contrast, defining "quality" for adolescents, adults, or aging populations can be more subjective and variable. What constitutes meaningful progress or enrichment later in life is harder to quantify, and normative benchmarks are less universal.

Key Indicator Methodology itself is also relatively rare in adult care regulatory systems. This may be due in part to the historical concentration of quality indicator research and application within child care, as well as the more mature and standardized nature of child care regulation compared to systems serving adults. As a result, further research and development will likely be needed to adapt and expand the key indicator framework for adult populations. This may include the creation of tiered, provider-level quality assessment tools that go beyond individual outcomes to measure systemic performance.

If such tools can be developed and validated, the integration of quality indicators into adult care licensing systems could mirror the transformative impact seen in child care, enabling targeted monitoring, data-informed licensing decisions, and a stronger link between compliance and meaningful quality outcomes.

About the Author:

Dr Richard Fiene, a research psychologist, has spent his professional career in improving the quality of child care in various states, nationally, and internationally. He has done extensive research and

publishing on the key components in improving child care quality through an Early Childhood Program Quality Indicator Model (ECPQIM) of training, technical assistance, quality rating & improvement systems, professional development, mentoring/coaching, regulatory science, licensing, risk assessment, differential program monitoring, key indicators, and accreditation. His research has also made significant contributions in regulatory science related to measurement and monitoring systems, such as instrument-based program monitoring, differential monitoring, key indicator methodology for compliance and quality, and risk assessment methodology. In prevention science, his research has led to the identification of key Regulatory indicators that keep children healthy and safe while in out of home child care settings.

Dr Fiene is a Professor of Psychology (ret) (Penn State University) and founding director of the Capital Area Early Childhood Research and Training Institute. He is presently a Research Psychologist and Regulatory Prevention Scientist for the Research Institute for Key Indicators, an affiliated data laboratory with the Edna Bennett Pierce Prevention Research Center at the Pennsylvania State University.

Dr Fiene is regarded as a leading international researcher/scholar on human services licensing measurement and differential monitoring systems. His regulatory compliance law of diminishing returns has altered human services regulatory science and licensing measurement dramatically in thinking about how best to monitor and assess licensing rules and regulations through targeted and abbreviated inspections. The theory has also led to the issuing of human service licenses based on substantial regulatory compliance with all rules rather than full 100% regulatory compliance with all rules. This was a basic licensing and public policy paradigm shift which has impacted on regulatory administration.

His research has led to the following developments: identification of herding behavior of two year olds, spatial acquisition device in young children & four states of space, national early care and education quality indicators, mathematical model (Contact Hours) for determining adult child ratio compliance, solution to the trilemma (quality, affordability, and accessibility) in child care delivery services, Stepping Stones to Caring for Our Children, NECPA: National Early Childhood Program Accreditation, online coaching as a targeted and individualized learning platform, validation framework for early childhood licensing systems and quality rating & improvement systems, an Early Childhood Program Quality Improvement & Indicator Model for better public policy decision making, Caring for Our Children Basics, Abbreviated Program Monitoring Inspections, Validation Framework for Licensing, Generic Key Indicator Rules, Regulatory Compliance Scoring Scale, Regal Metrics, and has led to the development of statistical techniques for dealing with highly skewed, non-parametric data distributions in human services licensing and regulatory systems, such as data dichotomization.

Dr Fiene had a long career in academia and governmental service. He was a research psychologist and regulatory scientist during his tenure with the Commonwealth of Pennsylvania's Office of Children, Youth, and Families and the Office of Licensing and Regulatory Administration where he was the research director for both offices. In academia, he was a professor of psychology and human development at both the University of North Carolina and the Pennsylvania State University. At Penn State Harrisburg he was Department Head for both the psychology and human development programs during his tenure at the university.

At the national and international levels, Dr Fiene has been a senior research consultant to the National Association for Regulatory Administration, the Federal Office of Child Care, the Administration for Children and Families, and the Federal Department of Health and Human Services. His research has been disseminated in all 50 states and over 120 countries. In 2019, he was elected to the Early Childhood Exchange Leadership Initiative. He received the 2020 Distinguished Career Award from the Pennsylvania Association for the Education of Young Children. In 2023, his Key Indicator methodology for quality indicators received a Recognized Project of the Child Impact Initiative of the World Forum Foundation. Dr Fiene remains active in the regulatory prevention science and early childhood fields through the Edna Bennett Pierce Prevention Research Center at Penn State where he remains an affiliated faculty and a senior research psychologist. He has been a member of the American Psychological Society.

Contributors

A special thank you to Ronald Melusky and Dr. Sonya Stevens for their contributions to this white paper.

Ronald Melusky is a regulatory professional with over twenty years of experience in regulatory administration and oversight. His work focuses on advancing systems of differential monitoring and strengthening the integrity of human services through effective policy development and implementation.

Dr. Sonya Stevens serves as NARA's Project Manager and brings a wealth of experience to the role. Her previous positions have encompassed licensing consultation, administration of statewide practice improvement initiatives, management of research and methodologies, and analytics concerning child care and foster care licensing. A passionate advocate for data-driven policies, Dr. Stevens is dedicated to fostering robust collaborations and partnerships that enhance the quality of services within human care regulation.

References and Future Readings:

American Academy of Pediatrics and American Public Health Association (1992). *Caring for Our Children*, AAP & APHA, Washington, D.C.

Fiene, R. & Nixon, M. (1985). Instrument based program monitoring and the indicator checklist for child care, *Child Care Quarterly, 14(3),* 198-214.

Fiene & McDonald (1987). *Instrument based program monitoring for child welfare*, Portland, Maine: University of Southern Maine.

Fiene (1995). *National early childhood program accreditation standards*. Atlanta, Georgia: National Early Childhood Program Accreditation Commission.

Fiene (1996). Using a statistical-indicator methodology for accreditation, in *NAEYC Accreditation: A Decade of Learning and the Years Ahead*, S. Bredekamp & B. Willer, editors, Washington, D.C.: National Association for the Education of Young Children.

Fiene (2002). Improving child care quality through an infant caregiver mentoring project, *Child and Youth Care Forum, 31*(2), 75-83.

Fiene, R. and Kroh, K.E. (2016). Human Services Licensing Measurement, Regulatory Compliance, and Program Monitoring Systems Chapter 11, part of *NARA Licensing Curriculum*, Research Institute for Key Indicators & the National Association for Regulatory Administration: Minneapolis, MN, USA.

Fiene, R. (2019). A treatise on regulatory compliance. *Journal of Regulatory Science* 7:1–3.

Fiene, R. (2022). Regulatory compliance monitoring paradigms and the relationship of regulatory compliance/licensing with program quality: A policy commentary. *Journal of Regulatory Science,* 10:1–7.

Fiene, R. (2024). *Regulatory compliance and monitoring systems ebook text*. National Association for Regulatory Administration licensing curriculum chapter on licensing measurement and program monitoring systems. Published online www.naralicensing.org/key-indicators.

Fiene, R. (2025a). The Child Care and Early Education Heart Monitor, submitted for publication in the *International Critical Childhood Policy Journal.*

Fiene, R. (2025b). The Development of a Regulatory Compliance Scale, *Scholarly Community Encyclopedia,* June, 2025, 1-18.

Fiene, R. (2025c). Finding the Rules that Work: An emerging paradigm promised to close the gap between regulatory compliance scores and quality of childcare services. *American Scientist, 113,* 16–21.

Fiene, R. (2025d). Potential Solution to the Child Care Trilemma Revisited: Finding the “Right Rules” - The Holy Grail of Early Care and Education, *Exchange*, Summer 2025, pps 64-66.

Fiene, R. (2025e). The Uncertainty–Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies. *Knowledge, 5(2), 8.* https://doi.org/10.3390/knowledge5020008

Freer, R. & Fiene, R. (2023). Regulatory Compliance and Quality Programs: Constraints and Opportunities for Integration. *Journal of Regulatory Science, 11(1)*, 1-10. https://doi.org/10.21423/JRS.REGSCI.111264.

Harms, T., Clifford, R., & Cryer, D. (2012). *Early Childhood Environmental Rating Scale*. Columbia University Press, New York, New York.

Kontos, S. & Fiene, R. (1987). Child care quality, compliance with regulations, and children's development: The Pennsylvania Study, in *Quality in Child Care: What Does Research Tell Us?*, Phillips, editor, Washington, D.C.: National Association for the Education of Young Children.

Lahti, M.; Elicker, J.; Zellman, G.; Fiene, R. (2014). Approaches to validating child care quality rating and improvement systems (QRIS): Results from two states with similar QRIS type designs. *Early Childhood Research Quarterly, 30*, 280–290.

Maxwell, K.; Miranda, B.; Harris, P.; Johnson, N.; Padilla, C.; Banghart, P.; Verhoye, A.; Early, D.M. (2024). *Child care and early education licensing literature review.* In OPRE Report #2024-199; Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services: Washington, DC, USA.

Miranda, B.; Ekyalongo, Y.; Franchett, A.; Maxwell, K. (2022). *Monitoring Practices Used in Child Care and Early Education Licensing.* In OPRE Report #2022-137; Office of Planning, Research, and Evaluation, Administration for Children and Families, Department of Health and Human Service: Washington, DC, USA.

Morgan, G. (1979). Methods for Achieving Quality Child Care, *Young Children, Volume 34, Number 6,* pps 22-27.

National Institute of Child Health and Human Development Early Child Care Research Network and Duncan, G.J. (2003), Modeling the Impacts of Child Care Quality on Children's Preschool Cognitive Development. Child Development, 74: 1454-1475. https://doi.org/10.1111/1467-8624.00617

Payne, A.L. (2011). *Strong Licensing: The Foundation for a Quality Early Care and Education System*, National Association for Regulatory Administration: Minneapolis, MN, USA.

Peisner-Feinberg, E.S., Burchinal, M.R., Clifford, R.M., Culkin, M.L., Howes, C., Kagan, S.L. and Yazejian, N. (2001), The Relation of Preschool Child-Care Quality to Children's Cognitive and Social Developmental Trajectories through Second Grade. *Child Development, 72*: 1534-1553. https://doi.org/10.1111/1467-8624.00364

Stevens, Fiene, Blevins, Salzer (2020). Identifying Predictive Indicators: The State of Washington Foster Care Home Study, *Journal of Children and Youth Services*.

Trivedi, P.A. (2015). *Innovation in Monitoring in Early Care and Education: Options for States*, Office of the Assistant Secretary for Planning and Evaluation, US Department of Health and Human Services: Washington, DC, USA.

Zellman, G. and Fiene, R. (2012). *Validation of quality rating and improvement systems for early care and education and school-age care.* In OPRE and Child Trends, Office of Planning Research and Evaluation Administration for Children and Families, U.S. Department of Health and Human Services: Washington, DC, USA.

Development of a Regulatory Compliance Scale

Subjects: Psychology

Contributor: Richard Fiene

A newly proposed Regulatory Compliance Scale is being suggested for consideration in regulatory science and human services regulatory administration and the licensing and review of programs and facilities. The article presents the parameters of the new scale along with research that has been done to date in using the proposed scale.

regulatory science | regulatory compliance | licensing measurement | human services regulatory administration

Development of a Regulatory Compliance Scale

Richard Fiene PhD

Penn State's Edna Bennett Pierce Prevention Research Center

April 2025

The purpose of this paper is to provide an alternate paradigm for regulatory compliance measurement in moving from a nominal to an ordinal scale measurement strategy and to introduce a new licensing/regulatory compliance metric: the Regulatory Compliance Scale. Regulatory compliance measurement is dominated by a nominal scale measurement system in which rules are either in compliance or out of compliance. There are no gradients for measurement within the present licensing measurement paradigm. It is very absolute. Either a rule is in full compliance to the letter of the law or the essence of the regulation or it is not. An alternate paradigm borrowing from accreditation and other program quality systems is to establish an ordinal scale measurement system which takes various gradients of compliance into account. With this alternate paradigm, it offers an opportunity to begin to introduce a quality element into the measurement schema. It also allows us to take into consideration both risk and prevalence data which are important in rank ordering specific rules.

So how would this look from a licensing decision making vantage point. Presently, in licensing measurement, licensing decisions are made at the rule level in which each rule is either in or out of compliance in the prevailing paradigm. Licensing summaries with corrective actions are generated from the regulatory compliance review. It is a nominal measurement system based upon Yes/No responses. The alternate measurement paradigm I am suggesting in this paper is one that is more ordinal in nature where we expand the Yes/No response to include gradients of the particular rule. In the next paragraph, I provide an example of a rule that could be measured in moving from a nominal to ordinal scale measurement schema.

Rather than only measuring a rule in an all or none fashion, this alternate paradigm provides a more relative mode of measurement at an ordinal level. For example, with a professional development or training rule in a particular state which requires, let's say, 6 hours of training for each staff person. Rather than having this only be 6 hours in compliance and anything less than this is out of compliance, let's have this rule be on a relative gradient in which any amount of hours above the 6 hours falls into a program quality level and anything less than the 6 hours falls out of compliance but at a more severe level depending on how far below the 6 hours and how many staff do not meet the requirement (prevalence). Also throw in a specific weight which adds in a risk factor, and we have a paradigm that is more relative rather than absolute in nature.

From a math modeling perspective, the 1 or 0 format for a Yes or No response becomes -2, -1, 0, +1, +2 format. This is more similar to what is used in accreditation systems where 0 equals Compliance and -1 and -2 equals various levels of Non-Compliance in terms of severity and/or prevalence. The +1 and +2 levels equal value added to the Compliance level by introducing a Quality Indicator. This new formatting builds upon the compliance vs non-compliance dichotomy (C/NC) but now adds a quality indicator (QI) element. By adding this quality element, we may be able to eliminate or at least lessen the non-linear relationship between regulatory compliance with rules and program quality scores as measured by the Environmental Rating Scales (ERS) and CLASS which is the essence of the Theory of Regulatory Compliance (TRC). It could potentially make this a more linear relationship by not having the data as skewed as it has been in the past.

By employing this alternate paradigm, it is a first demonstration of the use of the Key Indicator Methodology in both licensing and quality domains. The Key Indicator Methodology has been utilized a great deal in licensing but in few instances in the program quality domain. For example, over the past five years, I have worked with approximately 10 states in designing Licensing Key Indicators but only one state with Quality Key Indicators from their QRIS – Quality Rating and Improvement System. This new paradigm would combine the use in both. It also takes advantage of the full ECPQI2M – Early Childhood Program Quality Improvement and Indicator Model by blending regulatory compliance with program quality standards.

A major implication in moving from a nominal to an ordinal regulatory compliance measurement system is that it presents the possibility of combining licensing and quality rating and improvement systems into one system via the Key Indicator Methodology. By having licensing indicators and now quality indicators that could both be measured by licensing inspectors, there would be no need to have two separate systems but rather one that applies to everyone and becomes mandated rather than voluntary. It could help to balance both effectiveness and efficiency

by only including those standards and rules that statistically predict regulatory compliance and quality and balancing risk assessment by adding high risk rules.

I will continue to develop this scale measurement paradigm shift in future papers but wanted to get this idea out to the regulatory administration field for consideration and debate. This will be a very controversial proposal since state regulatory agencies have spent a great deal of resources on developing free standing QRIS which build upon licensing systems. This alternate paradigm builds off the Theory of Regulatory Compliance's key element of relative vs absolute measurement and linear vs non-linear relationships (Fiene, 2022). Look for additional information about this on RIKI Institute Blog - https://rikinstitute.com/blog/.

Introduction to the Regulatory Compliance Scale

The theory of regulatory compliance has been proven in multiple studies over the past four decades and has been utilized extensively in the creation of differential monitoring and its spin off methodologies of risk assessment and key indicators (Fiene, 2025). In fact, differential monitoring would not have been possible without the theory of regulatory compliance because the paradigm which it replaced, one of one-size-fits-all monitoring or uniform monitoring would have predominated. However, with the theory of regulatory compliance which introduced the importance of substantial regulatory compliance and the search for the right rules/regulations that made a difference in client's lives, rather than emphasizing more or less regulations or rules.

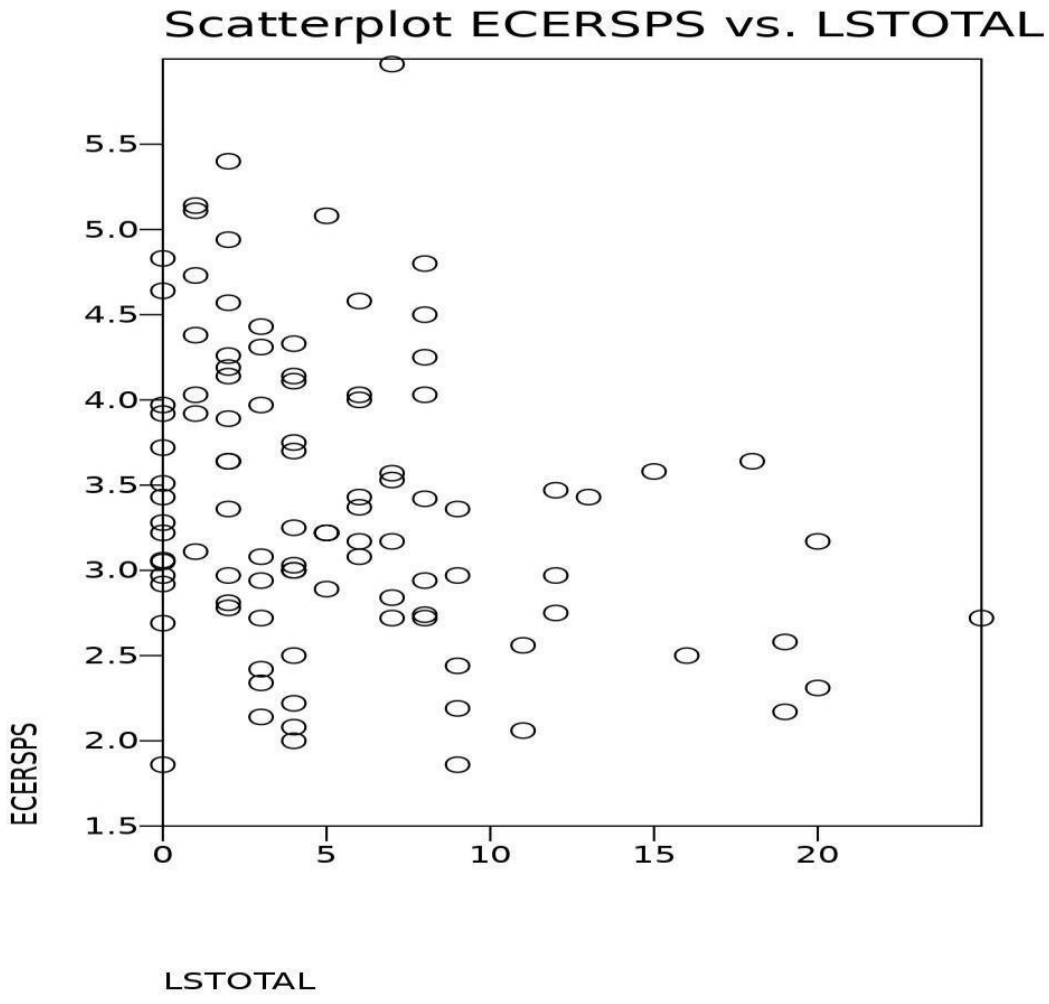
The theory of regulatory compliance has another application when it comes to regulatory compliance measurement in helping to move the licensing field from a nominal based measurement strategy to one of ordinal based measurement. The new measurement strategy is the Regulatory Compliance Scale (RCS) and it is depicted in the following table.

RCS	*Compliance*	*Risk*	*Model*	*Model*
Scale	***Level***	***Level***	***Violations***	***Weights***
7 = A	**Full**	**None**	**0**	**0**
5 = B	**Substantial**	**Low**	**1-3**	**1-3**
3 = C	**Medium**	**Medium**	**4-9**	**4-6**
1 = D	**Low**	**High**	**10+**	**7+**

The above table needs some explanation. The first column is the proposed ordinal scale similar to other scales utilized in the program quality measurement research literature on a 1 – 7 Likert Scale where 7 = Full Regulatory Compliance, 5 = Substantial Regulatory Compliance, 3 = Medium Regulatory Compliance, and 1 = Low Regulatory Compliance. It could also be thought of as an Alpha Scale of A – D as well. The next column has the compliance levels that run from full 100% regulatory compliance to low regulatory compliance. The third column depicts the risk level from none to high which corresponds with the compliance levels. The next two columns depict two models, one unweighted and one in which the rules are weighted with corresponding weights. These models are based upon the two prevailing approaches to rank ordering rules or regulations in the research literature.

The following figures will depict how the scale was conceived based upon empirical evidence in the various studies supporting the theory of regulatory compliance.

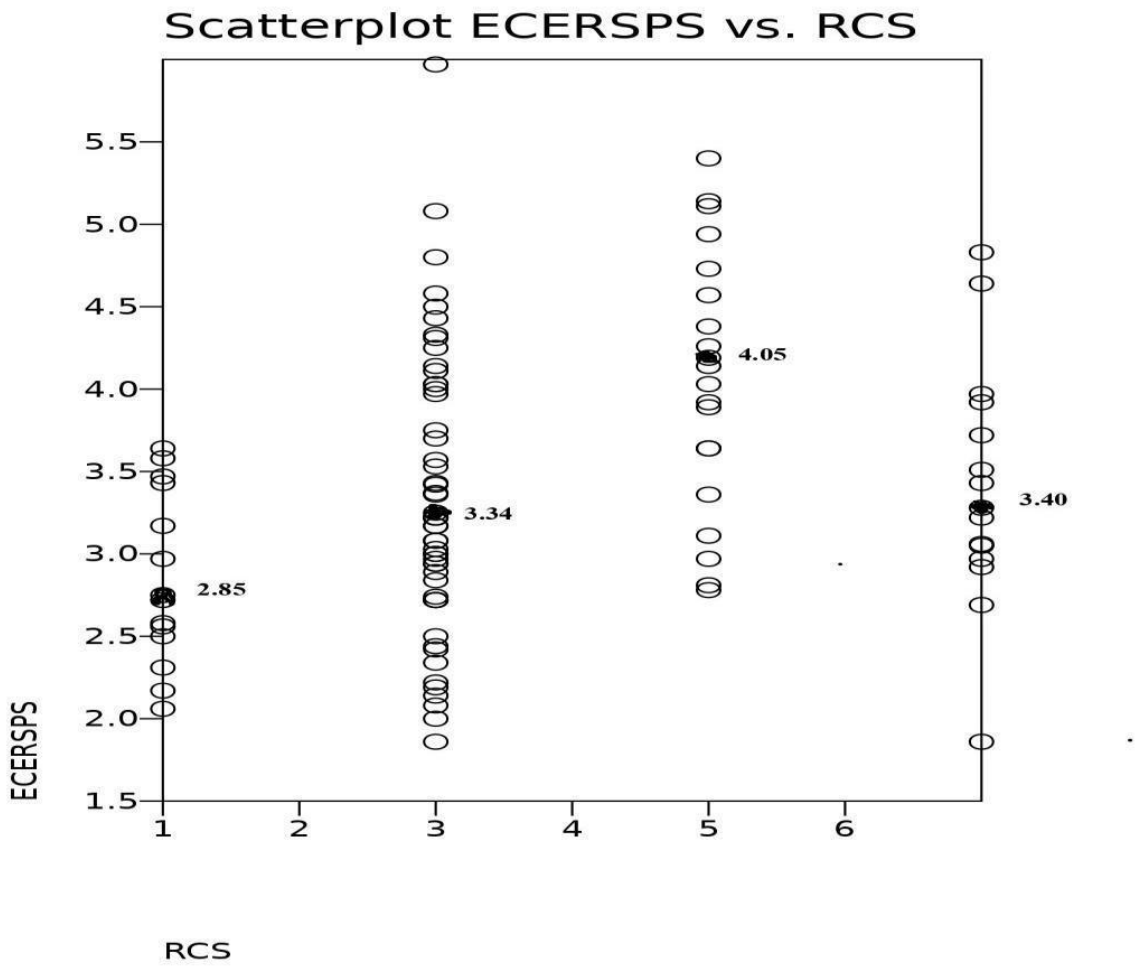
The first figure shows the actual individual violation data of the programs compared to their corresponding ECERS scores. There is not a significant relationship between the two as depicted in the graphic.



The following figure below depicts what occurs when the individual violation data are grouped according to the theory of regulatory compliance in which a substantial compliance category is introduced, and the data are moved from a nominally based metric to an ordinally based metric of full, substantial, medium, and low regulatory compliance categories. This grouping more clearly reflects the theory of regulatory compliance. It also clearly demonstrates the ceiling effect which is an outcome of the theory of regulatory compliance in which substantial and full regulatory compliance levels are basically equivalent when quality is taken into account. Or at the extreme level which is depicted here where full regulatory compliance quality scores are actually lower than the substantial regulatory compliance quality scores. A footnote about the figures and the scaling: the scales for the first figure are on a lower to higher progression but the higher LSTOTAL represents higher non-compliance where the second figure is also based upon lower to higher but the higher scores represent increased quality and increased regulatory compliance.

So, in reading the change from left to right, these two figures are reversed images of each other. This is just a quirk of the scaling and not a mistake in the plotting of data.

The RCS has been pilot tested in both the unweighted and weighted models and based upon these studies it appears to be more effective in distinguishing quality amongst the various categories rather than utilizing violation count data. This would be a significant improvement when it comes to licensing measurement. Of course, additional replication studies need to be completed before it would be recommended as a new Scale to be used for making licensing decisions.



The above figure is dramatically different than the prevailing paradigm which predicts a linear relationship between regulatory compliance and quality which is the paradigm of a uniform monitoring approach. The above results clearly indicate a reconsideration with the introduction of substantial regulatory compliance as an important contributor to overall quality if not the most important contributor to quality. As stated above, these findings have been replicated in several studies conducted over the past several decades.

This would be a major paradigm shift in moving from individual violation data counts to an ordinal scale metric but it does warrant additional research. The problem with individual violation data is that it doesn't take into account the relative risk of the individual rule which could place clients at increased risk of morbidity or mortality. Risk assessment has worked really well when coupled with key indicators in the differential monitoring approach and it appears to be an asset in the development of a Regulatory Compliance Scale (RCS).

Regulatory Compliance Scale Studies

The Regulatory Compliance Scale (RCS) was introduced several years ago and has been used in a couple of validation studies for differential monitoring and regulatory compliance's ceiling effect phenomenon. RCS buckets or thresholds were statistically generated based upon these studies, but it is time to validate those buckets and thresholds to determine if they are really the best model in creating a regulatory compliance scale. Since proposing the RCS, there has been a great deal of interest from jurisdictions in particular from Asian and African nations. Additional statistically based trials were conducted, and this brief report is the compilation of those trials over the past year.

The data used are from several jurisdictions that are part of the international database maintained at the Research Institute for Key Indicators Data Laboratory at Penn State University focusing on program quality scores and rule violation frequency data. These data from the respective databases were recoded into various thresholds to determine the best model. The jurisdictions were all licensing agencies in the US and Canada geographically dispersed where both regulatory compliance and program quality data was obtained from a sample of early care and education programs.

Methodology

The following methodology was used starting with the original RCS buckets/thresholds of Full, Substantial, Medium, and Low regulatory compliance:

RCS Models used for analyses

RCS				**Models**			
		Original	***1***	***2***	***3***	***4***	***5***
	Full	100	100	100	100	100	100
Scaling	***Substantial***	99-98	99-97	99-97	99-98	99-98	99-97
	Medium	97-90	96-90	96-93	97-95	97-85	96-85
	Low	89>	89>	92>	94>	84>	84>

Five alternate models were used to compare the results to the original RCS. The numbers indicate the number of violations subtract from a perfect score of 100. Full regulatory compliance indicates no violations and a score of 100 on the scale. The next bucket of 99-98 indicates that there were 1 or 2 regulatory compliance violations which resulted in a 99-98 score on the scale. This logic continues with each of the models.

The scale score was determined in the following manner: Full Regulatory Compliance = 7; Substantial Regulatory Compliance = 5; Medium Regulatory Compliance = 3; and Low Regulatory Compliance =1. This rubric is how the original RCS scaling was done on a Likert type scale similar to other ECE program quality scales, such as the Environmental Rating Scales.

Results

The following results are correlations amongst the respective RCS Models from Table above compared to the respective jurisdictions program quality tool (Quality1-3): ERS or CLASS Tools.

RCS Model Results compared to Quality Scales

RCS results	Models	Quality1	Quality2	Quality3
Jurisdiction1	**RCS0**	**.26***	**.39***	**.39***
	RCS3	.21	.32*	.33*
	RCS5	.20	.36*	.33*
Jurisdiction2	**RCS0**	**.76****	**.46****	**---**
	RCS3	.12	-.07	---
	RCS5	.18	-.02	---
	RCSF1	.55**	.29*	---
	RCSF2	.63**	.34	---
Jurisdiction3	RCS0	.19	.18	.16
	RCS3	**.21**	**.21**	**.15**
	RCS5	.18	.16	.07
	RCSF1	.17	.17	.10
	RCSF2	.18	.18	.19
Jurisdiction4	RCS0	.24*	---	---
	RCS3	.28*	---	---
	RCS5	**.30***	**---**	**---**
	RCSF1	.21	---	---
	RCSF2	.29*	---	---
Jurisdiction5	RCS0	.06	-.02	.07
	RCS3	.06	-.01	.05

	RCS5	**.08**	**.00**	**.09**
	RCSF1	.00	-.03	.05
	RCSF2	.05	-.03	.05

*Statistically significant .05 level;

**Statistically significant .01 level.

In the above table starting under Jurisdiction2, two new models were introduced based upon the

Fibonacci Sequence (Fibonacci1 = RCSF1; Fibonacci2 = RCSF2) and their model structure is in the following Table. The reason for doing this is that the Fibonacci Sequence introduces additional variation into the scaling process.

RCS Fibonacci Models

RCS Fibonacci			**Models**	
		Original	***Fibonacci1***	***Fibonacci2***
	Full	100	100	100
Scaling	***Substantial***	99-98	40	90
	Medium	97-90	20	20
	Low	89>	13	13

A second series of analyses were completed in comparing the RCS models with program quality (Quality1) by running ANOVAs with the RCS models as the independent variable and program quality as the dependent variable. The reason for doing this was the nature of the data distribution in which there was a ceiling effect phenomenon identified which would have had an impact on the correlations in table above. All results are significant at $p < .05$ level with the exception of Jurisdiction2.

ANOVAs Comparing the RCS Models with Program Quality

Jurisdictions	Model	Level 1	Level 3	Level 5	Level 7
Jurisdiction1	**RCS0**	**2.85**	**3.34**	**4.05**	**3.40**
	RCS3	3.24	3.23	4.05	3.40
	RCS5	2.73	3.32	3.77	3.40
Jurisdiction2	RCS0	4.81	4.31	4.80	4.10
	RCS3	**4.59**	**4.25**	**4.80**	**4.10**
	RCS5	---	4.26	4.64	4.10
Jurisdiction3	RCS0	4.59	4.68	4.86	4.87
	RCS3	**4.38**	**4.67**	**4.83**	**4.87**
	RCS5	4.38	4.83	4.83	4.87
Jurisdiction4	RCS0	37.81	37.01	44.28	41.96
	RCS3	36.57	38.60	44.28	41.96
	RCS5	**33.46**	**36.53**	**43.10**	**41.96**
Jurisdiction5	RCS0	3.93	4.17	4.28	4.07
	RCS3	4.02	4.24	4.28	4.07
	RCS5	**3.75**	**4.13**	**4.26**	**4.07**

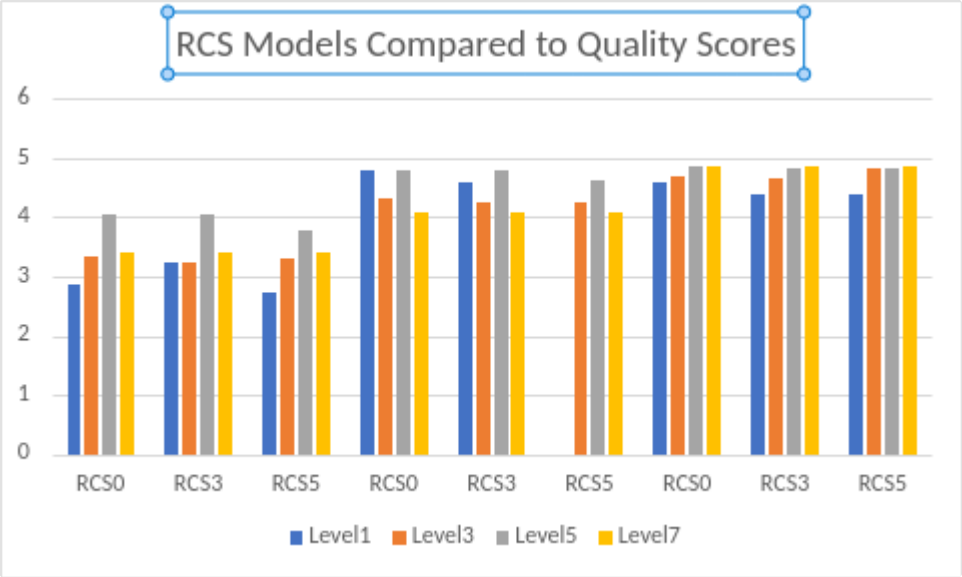
Insights

Based upon the above results, it appears that the original RCS model proposed in 2021 is still the best model to be used, although the Fibonacci Sequence model is a close second in some of the jurisdictions. This model will need further exploration in determining its efficacy as a replacement or enhancement to the original RCS Model.

The bottom line is that the original RCS Model is as good as any and no other model is consistently better than all the rest. The RCS Model does have a slight edge over Regulatory Compliance Violation RCV frequency counts in some jurisdictions but not in others. It is much easier to interpret the relationship between quality and the RCS

models than it is to interpret the results from the quality scores and the RCV data distribution. So, the recommendation would be for licensing agencies to think about using this new scaling technique in one of its model formats to determine its efficacy. Pairing up RCS and RCV data side by side by licensing agencies would be important studies to determine which approach is the better approach.

The below graphic depicts the relationship between the RCS Models (0, 3, 5) when compared to the quality scores (1-6) clearly showing the ceiling effect and diminishing returns effect phenomenon so typical of regulatory compliance data when compared to program quality. These graphs are from the first three jurisdictions (1, 2, 3) from the above tables.



Additional Analyses Comparing the 11 Studies

This section provides the results from 11 studies from 10 states and Canadian Provinces in which the proposed new Regulatory Compliance Scale (RCS) was utilized as a byproduct of a differential monitoring implementation or validation study. These studies were undertaken over a decade long period (2013-2023).

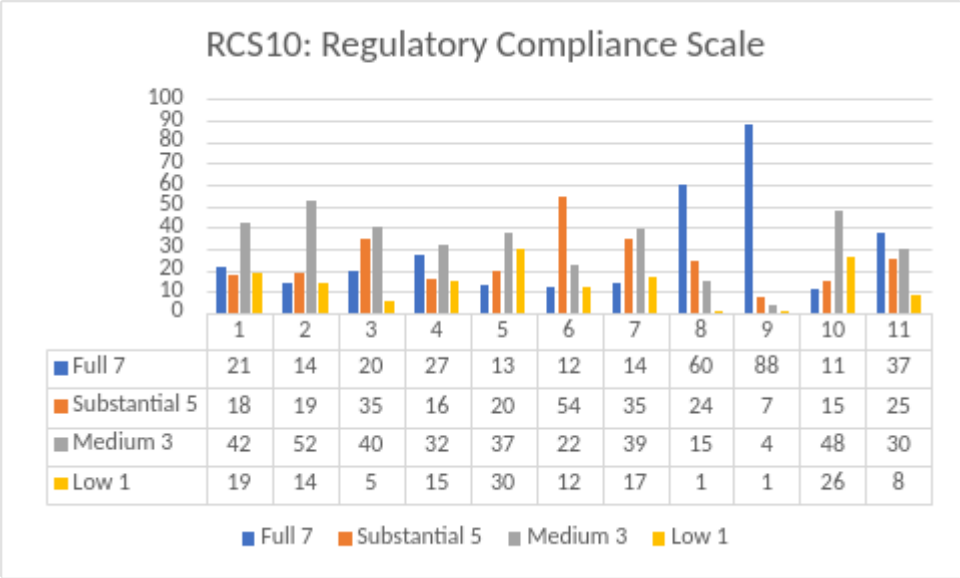
The RCS was based upon the following rubric: Full Regulatory Compliance (100%) or no violations = 7; Substantial Regulatory Compliance (99-98) or 1-2 violations = 5; Medium Regulatory Compliance (97-90) or 3-10 violations =3; and Low Regulatory Compliance (89 or less) or 11 or more violations = 1.

These are the results from these 10 jurisdictions which are presented in the following Table (all results are presented as percentages of programs that fell into the scaling 1-7). Under the Studies, the number of the specific

study is provided, followed by the sample size, followed by if it is in the USA (US) or Canada (CA).

RCS Scale			**RCS Scaling**		
Studies	**7=Full**	**5=Substantial**	**3=Medium**	**1=Low**	**Comments**
1-403-US	21%	18%	***42%***	19%	***High Med NC***
2-104-US	14%	19%	***52%***	14%	***High Med NC***
3-422-US	***20%***	***35%***	***40%***	***5%***	***OK***
4-219-CA	***27%***	***16%***	***32%***	***15%***	***OK***
5-60-CA	13%	20%	***37%***	***30%***	***High NC/Low C***
6-585-US	***12%***	***54%***	***22%***	***12%***	***OK***
7-255-US	***14%***	***35%***	***39%***	***17%***	***OK***
8-1399-US	***60%***	24%	15%	***1%***	***Low NC/High C***
9-2116-US	***88%***	7%	4%	***1%***	***Low NC/High C***
10-482-US	11%	15%	*48%*	*26%*	***High NC/Low C***
11-3070-US	***37%***	***25%***	***30%***	***8%***	***OK***

In looking at the results, it is preferable to have most of the programs at either a full or substantial regulatory compliance level (7 or 5) and to have fewer programs at the medium or low regulatory compliance level (3 or 1). But in those jurisdictions where there are higher percentages of programs at the medium or low levels of regulatory compliance, it could be that their enforcement of rules and regulations is more stringent. This potential result needs further investigation to get to the root cause of these differences because there is a good deal of variation across the jurisdictions as is evident from the above table.



	1	2	3	4	5	6	7	8	9	10	11
Full 7	21	14	20	27	13	12	14	60	88	11	37
Substantial 5	18	19	35	16	20	54	35	24	7	15	25
Medium 3	42	52	40	32	37	22	39	15	4	48	30
Low 1	19	14	5	15	30	12	17	1	1	26	8

Based upon the above studies and results, the regulatory compliance scale (Fiene, 2022) which appears from recent studies to be a better metric in measuring regulatory compliance than just counting the number of violations that a program has related to their respective rules, regulations, or standards. So how does the regulatory compliance scale work? It essentially puts violations into buckets of regulatory compliance as follows: full compliance (100%) or no violations; substantial compliance (99-98%) or 1-2 violations; mediocre compliance (97-90%) or 3-9 violations; and lastly low/non-optimal compliance (89% or lower) or 10+ violations. Why buckets, because logically it works, it is the way we think about regulatory compliance. It is a discrete rather than continuous metric and logically fits into these four categories. This is based upon 50 years of research into regulatory compliance data distributions and when the data are moved from frequency counts of violation data into these buckets/categories, the math works very well in identifying the better performing programs.

Regulatory Compliance Scale Extensions

Depicted below is a regulatory compliance grid model showing the relationship between regulatory compliance (RC) and program quality (PQ).

An explanation of the below chart will demonstrate how regulatory compliance and program quality in human service facilities interact. The horizontal blue axis depicts the various levels of regulatory compliance while the vertical green axis depicts the various levels of program quality of facilities. It ranges from 1-5 or low to high for each axis. The red "**X's"** represents the relationship that has been identified in the research literature based upon the theory of regulatory compliance in which there is either a plateau effect or a downturn in quality as regulatory compliance increases. The one italicized "***X***" is an outlier that has also been identified in the research literature in which some (it does not happen often) low compliant programs really are at a high-quality level.

It is proposed in order to mitigate the plateau effect with regulatory compliance and program quality standards because regulatory compliance data distributions are severely skewed which means that many programs that have questionable quality are being included in the full (100%) compliance domain. When regulatory compliance

standards are increased in their quality components this will lead to a higher level of overall quality as depicted in the "**XX**" cell all the way on the lower right. It also helps to mi gate the severe skewness in the regulatory compliance data distribution. The data distribution does not approximate a normally distributed curve which is the case with the program quality data distribution.

Regulatory Compliance x Program Quality Grid Model

PQ/RC ->	1 Low	2 Med	3 Substantial	4 Full 100%	5QualityAdditions
1 Low	XXX				
2		XX			
3 Med			XX	XXX	
4			XX	X	
5 High	*X*				XX

By utilizing this model, it helps to deal more directly in taking a non-linear relationship and making it linear again when comparing regulatory compliance with program quality. This model provides a theoretical approach supporting what many state licensing administrators are thinking from a policy standpoint: add more quality to health and safety rules/regulations. This grid/matrix also depicts the three regulatory compliance models: Linear, Non-linear, and Stepped.

Here is another potential extension of the Regulatory Compliance Scale using the ECPQIM DB – Early Childhood Program Quality Improvement and Indicator Model Database, it is possible to propose developing and using a Regulatory Compliance Scoring System and Scale (RC3S). This new proposed RC3S could be used by state human service agencies to grade facilities as is done in the restaurant arena. Presently, in the human service field, licenses are issued with a Certificate of Compliance but generally it does not indicate what the regulatory compliance level is at. This new proposal would alleviate this problem by providing a scale for depicting the level of regulatory compliance.

The ECPQIM DB is an international database consisting of a myriad group of data sets drawn from around the USA and Canada. It has been in the making over 40 years as of this writing, so its stability and generalizability have been demonstrated. What follows is the chart depicting the RC3S.

Regulatory Compliance Scoring System and Scale (RC3S)

Color	Non-Compliance Level	Regulatory Compliance Level
Blue	0	Full Compliance
Green	1-2	Substantial Compliance
Yellow	3-6	Mid-Range Compliance
Orange	7-9	Low Compliance
Red	10-15+	Very Low Compliance

It is evident from the above chart that the color go from blue to red which indicates an increasing risk of non-compliance and a lower level of overall regulatory compliance, which is not a good thing in the licensing field. Non-compliance levels indicate the number of rules or regulations or standards that are not complied with. And lastly, the regulatory compliance level indicates the movement from full (100% regulatory compliance with all rules) to very low compliance with rules. These ranges for the scaling are based on 40 years of research in understanding and plotting the data distributions around the world related to regulatory compliance in the human services. These results have consistently appeared over this 4-decade time period and show no signs of changing at this point.

Regulatory Compliance Scaling for Decision Making

There is a lack of empirical demonstrations of regulatory compliance decision making. In the past, I have used the methodologies of key indicators, risk assessment and the resultant differential monitoring techniques of how often and what should be reviewed for decision making. What has not been addressed is decision making based upon comprehensive reviews when all regulations are assessed. This section addresses how empirical evidence taken from the past 40+ years of establishing and researching a national database for regulatory compliance can help lead us to a new scaling of regulatory compliance decision making.

In analyzing regulatory compliance data, it becomes perfectly clear that the data have very little variance and are terribly skewed in which the majority of programs are in either full or substantial compliance with all the respective regulations. Only a small handful of programs fall into the category of being in low compliance with all the regulations.

The proposed scaling has three major decision points attached to regulatory compliance scores. Either programs are in full or substantial compliance, in low compliance or somewhere in the middle. Full or substantial regulatory compliance is 100% or 99-98% in regulatory compliance. Low regulatory compliance is less than 90% and mid-regulatory compliance is between 97%-90%. These ranges may seem exceptionally large but based upon the

national database on regulatory compliance that I maintain at the Research Institute for Key Indicators (RIKILLC) these are the ranges that have formed over the past 40 years. These data ranges should not come as a surprise because we are talking about regulatory compliance with health and safety standards. These are not quality standards; these are basic protections for clients. The data are not normally distributed, not even close as is found in quality tools and standards.

What would a **Regulatory Compliance Decision-Making Scale** look like:

Data Level Decision_____

100-98% Full/Substantial License

97-90% Mid-Range Provisional

89% or less Low No-License

States/Provinces/Jurisdictions may want to adjust these levels, and the scaling based upon their actual data distribution. For example, I have found certain jurisdictions to have very unusually skewed data distributions which means that these ranges need to be ghten even more. If the data distribution is not as skewed as the above scale, then these ranges may need to be more forgiving.

This regulatory compliance decision making scale does not take into account if abbreviated methodologies are used, such as risk assessment or key indicator models that are used in a differential monitoring approach. The above scale is to be used if a jurisdiction decides not to use a differential monitoring approach and wants to measure regulatory compliance with all regulations and complete comprehensive reviews.

Conclusion

The Theory of Regulatory Compliance (Fiene, 2019) and bringing substantial compliance to the forefront of regulatory science has been written about a great deal. This paper builds upon these previous assertions and expands them into some practical applications that can be utilized within regulatory science as it relates to licensing measurement, regulatory compliance scaling, and monitoring systems paradigms. This paper has introduced the Regulatory Compliance Scale which is a departure in how best to measure regulatory compliance. This new scale along with the proposed Uncertainty-Certainty Matrix (Fiene, 2025b) provides a robust licensing measurement and program monitoring strategy. This paper provides the last piece of a differential monitoring approach that includes instrument-based program monitoring, key indicators, risk assessment, and the uncertainty-certainty matrix.

Regulatory Compliance has been always approached as an all or none phenomenon, whether a rule is in compliance, or it is not. There is no in-between or shades of gray or partial compliance. This worked when the prevailing paradigm was that full regulatory compliance and program quality were a linear relationship. This was the assumption but not empirically verified until the later 1970's-1980's. When this assumption was put to an empirical test, it did not hold up but rather a curvilinear relationship between regulatory compliance and program quality was discovered. This upset the prevailing paradigm and suggested we needed a new approach to addressing the relationship between regulatory compliance and program quality.

It became clear after these findings in the 1970's-80's and then in the 2010's when replication studies were completed that substantial regulatory compliance could not be ignored based upon this new theory of regulatory compliance in which substantial compliance acted as a "sweet spot" of best outcomes or results when comparing regulatory compliance and program quality scores. The nominal metric needed to be revised and more of an ordinal metric was to be its replacement. Because now it wasn't just being in or out of compliance, but it mattered which rules were in or out of compliance and how they were distributed. This revised application involved aggregate rules and does not apply to individual rule scoring. The studies completed between 1970 and 2010 involved aggregate rules and not individual rules. To determine if the nominal to ordinal metric needs to be revised still needs empirical data to back this change.

The introduction of substantial compliance into the regulatory compliance measurement strategy moved the field from an instrument-based program monitoring into a more differential monitoring approach. With differential monitoring this approach considered which rules and how often reviews should be done. Also, a new Regulatory Compliance Scale was proposed to take into account the importance of substantial compliance based upon the regulatory compliance theory of diminishing returns. As this Regulatory Compliance Scale has evolved within the licensing health and safety field it needs further revision in which program quality can be infused into the decision making related to individual rules. Remember that the original studies were concerned about rules in the aggregate and not individual rules. It has now become apparent that in dealing with the infusion of quality into rule formulation, a return to the individual rule approach makes the most sense.

The next iteration of the Regulatory Compliance Scale will contain the following categories: Exceeding full compliance, Full compliance, Substantial compliance, and Mediocre compliance to adjust for the infusion of the quality element. This differs slightly from the original aggregate rule Regulatory Compliance Scale where the categories were Full compliance, Substantial compliance, Mediocre compliance and Low compliance where only licensing health and safety elements were considered (see the Table below which depicts the regulatory compliance scales and program monitoring systems side by side).

Without the Theory of Regulatory Compliance, differential and integrative monitoring would not be needed because regulatory compliance would have had a linear relationship with program quality and full compliance would have been the ultimate goal. There would have been no need for targeted rule enforcement or reviews because all rules would have had an equal weight when it came to protecting clients and any individual rule would have predicted

overall compliance. But it “just ain’t so” as it is said. The need to make adjustments is brought about by the theory and it has not been the same ever since.

Regulatory Compliance Scales and Program Monitoring Systems

Scoring Level	**Individual Rule**		**Aggregate Rules**	**Individual Rule**
Scale	**Instrument based**	***Scale***	**Differential**	**Integrated**
7	Full Compliance	7	Full Compliance	Exceeds Compliance
-	---	5	Substantial	Full Compliance
-	---	3	Mediocre	Substantial
1	Out of Compliance	1	Low	Mediocre/Low

The above table attempts to summarize in tabular form the previous paragraphs in describing the relationship between program monitoring and licensing measurement scaling via a proposed regulatory compliance scale. As one can see this moves the paradigm from a nominal to an ordinal measurement rubric and depicts the differences in the measurement focus either at the individual rule or aggregate rules scoring levels. It also considers the significance of substantial compliance given the theory of regulatory compliance in which substantial compliance focus is a “*sweet spot*” phenomenon as identified in the regulatory science research literature. It is hoped that the regulatory science field takes these paradigm shifts into consideration in moving forward with building licensing decision making systems and how licenses are issued to facilities.

As a final footnote, keep in mind that the Theory of Regulatory Compliance applies to the relationship between regulatory compliance and program quality and does not apply to regulatory compliance in and of itself related to health and safety. When dealing with regulatory compliance, full compliance is the ultimate goal with individual rules and in determining which rules are predictive rules. It is the preferred methodology in order to eliminate false negatives and decreasing false positives in making licensing decisions related to regulatory compliance.

These above concepts all relate to the field of regulatory compliance and how to make informed decisions about licensing, particularly in the context of program monitoring. Here's how they connect:

Regulatory Compliance Scales:

These scales move away from a binary "compliant" or "non-compliant" approach to regulations. Instead, they acknowledge degrees of compliance, recognizing that minor deviations may not be as detrimental as major ones.

They provide a framework for evaluating the severity and frequency of non-compliance, allowing for more nuanced licensing decisions.

Instrument Based Program Monitoring (IBPM):

This is the traditional method of monitoring compliance, relying on standardized instruments and checklists to assess adherence to specific rules.

It's a comprehensive approach, but can be time-consuming and inflexible, potentially leading to over-regulation or missing important aspects of program quality.

Differential Monitoring (DM):

This approach takes into account the risk associated with different regulations, focusing monitoring efforts on areas with the highest potential for harm or non-compliance.

It allows for a more efficient use of resources and can be tailored to the specific needs of each program.

DM often utilizes Regulatory Compliance Scales to determine the severity of non-compliance and guide the level of monitoring needed.

Integrative Monitoring Systems (IMS):

These systems go beyond simply checking compliance and aim to assess the overall quality of a program.

They integrate data from various sources, including IBPM, DM, and other program-specific metrics, to provide a holistic picture of performance.

IMS can inform licensing decisions by considering not only compliance but also program effectiveness in achieving its goals.

Here's a simplified analogy to illustrate the relationships:

Think of regulations as traffic rules.

IBPM is like a police officer checking every car for every violation, regardless of severity.

DM is like a police officer focusing on patrolling areas with high accident rates or known reckless drivers.

Regulatory Compliance Scales are like different levels of fines based on the severity of the traffic violation.

IMS is like a traffic management system that collects data on accidents, traffic flow, and road conditions to optimize traffic flow and safety.

Relationships:

RCS forms the foundation for DM and IMS by providing a way to assess degrees of compliance.

IBPM provides data for RCS and can be incorporated (with adaptations) into DM and IMS.

DM builds on RCS and IBPM by differentiating the intensity of monitoring based on risk and compliance.

IMS is the most comprehensive approach, integrating RCS, IBPM, DM, and additional data sources for a deeper understanding of program performance.

Regulatory Compliance Scales can be used within any of the monitoring approaches to provide a more nuanced assessment of compliance.

IBPM can be a starting point for differential monitoring, providing data on rule compliance to inform risk assessments.

Differential monitoring can be integrated into an integrative monitoring system, along with other data sources, to provide a comprehensive picture of program performance.

Here are some additional points to consider:

The choice of the most appropriate approach will depend on the specific context, such as the type of program being regulated and the available resources.

Implementation of these alternative paradigms requires careful planning and training of regulators and program providers.

Ongoing research and evaluation are needed to refine these approaches and ensure their effectiveness.

These alternative paradigms offer a more flexible and effective approach to licensing decisionmaking compared to the traditional IBPM approach. They allow for a better understanding of program strengths and weaknesses, optimize resource allocation, and ultimately lead to better regulatory outcomes.

These concepts offer a shift from traditional "one-size-fits-all" compliance models to more flexible and nuanced approaches that consider risk, program quality, and degrees of compliance. This can lead to more efficient and effective regulatory systems that support program improvement while protecting public safety.

Ultimately, these concepts offer alternative paradigms for licensing decision-making, moving away from a rigid "one-size-fits-all" approach to a more nuanced and risk-based system that considers both compliance and program quality.

References:

Fiene, R. (2019). A treatise on Regulatory Compliance. ***Journal of Regulatory Science, Volume 7***, 2019. https://doi.org/10.21423/jrs-v07fiene

Fiene (2022a). Regulatory Compliance Monitoring Paradigms and the Relationship of Regulatory Compliance/Licensing with Program Quality: A Policy Commentary. (2022). ***Journal of Regulatory Science*, *10*(1). https://doi.org/10.21423/JRS-V10A239**

Fiene (2022b). Regulatory Compliance Scale, ***RIKINotes Blog***, January 2022.

Fiene (2023a). ***Licensing Measurement & Monitoring Systems***, Research Institute for Key Indicators, Elizabethtown, Pennsylvania.

Fiene (2023b). Ceiling Effect/Diminishing Returns, Regulatory Compliance Scale, and Quality Indicators Scale, ***Mendeley Data,*** doi: 10.17632/gc423hprcs.1

Fiene (2025a). Finding the Right Rules. ***American Scientist, Volume 113, 1***. pps 16-19.

Fiene (2025b). The Uncertainty-Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies, ***Knowledge Journal***, under review.

NARA (2023). ***Saskatchewan Differential Monitoring/Quality Indicators Scale Validation Study***, National Association for Regulatory Administration, Fredericksburg, Virginia.

Richard Fiene PhD, Senior Research Psychologist, Research Institute for Key Indicators (RIKILLC), Edna Bennett Pierce Prevention Research Center; Professor of Psychology (ret), Penn State University; Senior Research Consultant, National Association for Regulatory Administration (NARA). Email: Rfiene@RIKInstitute.com Website: http://RIKInstitute.com

---

knowledge

Study Protocol

The Uncertainty–Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies

Richard Fiene

MDPI

https://doi.org/10.3390/knowledge5020008

knowledge

MDPI

Study Protocol

The Uncertainty–Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies

Richard Fiene

Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University, University Park, PA 16802, USA; rfiene@rikinstitute.com

Abstract: This research article proposes the use of an uncertainty–certainty matrix (UCM) for licensing decision making in the human services, which is the decision to issue a license to operate. It is a proposed study protocol and conceptual framework; it is not an empirical study. It shows how the matrix can be used in rule decision making and how it clearly shows when decision making has gone awry when bias is introduced into the decision making. It is also proposed to be used to make decisions in differential monitoring and in validation and reliability studies. This proposal presents a potential blueprint on how the UCM can be used within human services licensing as a decision-making tool.

Keywords: decision making; uncertainty–certainty matrix; regulatory compliance; licensing; reliability and validation studies

Academic Editor: Cengız Kahraman

Received: 1 January 2025
Revised: 25 April 2025
Accepted: 27 April 2025
Published: 28 April 2025

Citation: Fiene, R. The Uncertainty–Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies. *Knowledge* **2025**, *5*, 8. https://doi.org/10.3390/knowledge5020008

1. Introduction

This research proposal takes the Contingency Table, which is a well-known metric in the statistical decision-making research literature [1], and refocuses it on regulatory science within the context of the definition of regulatory compliance and licensing measurement. It also deals with the policy implications of this particular metric. In this study protocol, it is proposed that the Uncertainty–Certainty Matrix (UCM) is a fundamental building block to licensing decision making from a measurement perspective. The Contingency Table, as demonstrated by a 2 × 2 matrix, is utilized in regulatory compliance and is the center piece for determining licensing key indicator rules [2], but it is also a core conceptual framework in licensing measurement and ultimately in program monitoring and reviews [3].

The reason for selecting this matrix is the nature of licensing data: it is binary or nominal in measurement. Either a rule/regulation is in compliance or out of compliance. Presently, most jurisdictions deal with regulatory compliance measurement in this nominal level or binary level. There is to be no gray area; this is a clear distinction in making a licensing decision about regulatory compliance. The UCM also takes the concept of Inter-Rater Reliability (IRR) a step further in introducing an uncertainty dimension that is very important in licensing decision making which is not as critical when calculating IRR. Inter-Rater Reliability is a real concern in the human services licensing field in that in many cases it is difficult for jurisdictions to maintain a high degree of consistency when comparing individual licensing inspectors to each other. Part of the problem is a fundamental measurement issue; it is hoped that the addition of the UCM will help to mitigate this problem [4]. Licensing measurement is dominated by nominal measurement: either a rule is in compliance or it is out of compliance. A proposal has been suggested in which an ordinal scale based upon licensing rule violations would be utilized called the

Regulatory Compliance Scale (RCS) [3]. This new RCS scale shows promise, but it needs additional validation studies in order to be used on a regular basis for making human services licensing decisions (2a).

The reason for even suggesting this matrix is the high level of dissatisfaction with the levels of reliability in the results of program monitoring reviews as suggested in the previous paragraph. If the dissatisfaction was not so pronounced, it would not be an issue, but with it being so high, the field of licensing needs to take a proactive role in determining the best possible way to deal with increasing inter-rater reliability among licensing inspectors. Hopefully, this organizational schema via the UCM Matrix will help to think through this process related to licensing measurement and monitoring systems. The author has been conducting regulatory compliance studies for the past 50 years and has determined that the validity and reliability of these studies needs a more robust model for making licensing decisions via more accurate measurements of regulatory compliance. This led to the creation and proposing of the UCM Metric [5–7].

Over the past 50 years, it has been well documented by the National Association for Regulatory Administration (NARA) how the licensing field has changed in moving from a one-size-fits-all licensing and monitoring approach to one of differential or targeted licensing and monitoring (https://www.naralicensing.org/key-indicators, accessed on 24 April 2025). NARA has led this transition in the human services licensing and regulatory administration field, which has produced a much more productive, effective, and efficient licensing inspection system. The UCM and RCS are the latest pieces in the puzzle to accomplishing this new licensing decision-making framework.

The key pieces to the UCM are the following: the decision (D) regarding regulatory compliance and actual state (S) of regulatory compliance. Regulatory Compliance of individual Rules: Plus (+) = In-compliance, or Minus (−) = Out of compliance. As such, the matrix can be built as follows (Table 1):

Table 1. Uncertainty–Certainty Matrix (UCM) Logic Model.

UCM Matrix Logic		**Decision (D) Regarding**	**Regulatory Compliance**
		(+) In Compliance	(−) Not In Compliance
Actual State (S) of Compliance	(+) In Compliance	Agreement	Disagreement
	(−) Not In Compliance	Disagreement	Agreement

The above UCM matrix demonstrates when agreement and disagreement occur, which establishes a level of certainty (Agreement Cells) or uncertainty (Disagreement Cells). In a perfect world, there would only be agreements and no disagreements between the decisions made about regulatory compliance and the actual state of regulatory compliance. However, from experience, this is not the case. This is based up reliability testing carried out in the human services licensing research field in which a decision is made regarding regulatory compliance with a specific rule or regulation, and then that is verified by a second licensing inspector observer who generally is considered the measurement standard.

Disagreements raise concerns in general, but the disagreements are of two types: false positives and false negatives. A false positive is when a decision is made that a rule/regulation is out of compliance when it is in compliance. This is not a good thing, but its twin disagreement is worse. With false negatives, it is decided that a rule/regulation is in compliance when it is out of compliance. False negatives need to be avoided because they place clients at extreme risk more so than a false positive. False positives should also be avoided, but it is more important to deal with the false negatives first before addressing the false positives.

2. Uncertainty–Certainty Matrix for Validation and Reliability Studies

This part of the research proposal is to explore the possibility of utilizing the Uncertainty–Certainty Matrix (UCM) as depicted in Table 1 in validation and reliability studies in licensing decision making. The UCM has been proposed for use in licensing decision making, but this would be an extension of this thinking to studies that involve validating licensing decisions, such as when key indicators/predictor rules are used in comparison with comprehensive reviews of rules [5] and in reliability studies to determine individual inspector bias in regulatory compliance [8,9].

The basic premise of the UCM is that individual decision making matches reality. When it comes to regulatory compliance decision making. a 2 × 2 matrix can be drawn with the possible outcomes as indicated in the following table (Table 2), which is based upon the logic of Table 1.

Table 2. Uncertainty–Certainty Matrix (UCM) Logic Model applied to Validation Studies.

UCM Matrix Logic	For Validation Studies	Decision Regarding	Regulatory Compliance
		(+) In Compliance	(−) Not In Compliance
Actual State of Compliance	(+) In Compliance	Agreement (++)	Disagreement (+−)
	(−) Not In Compliance	Disagreement (−+)	Agreement (−−)

In using this table, the hope is that the decision regarding regulatory compliance matches the actual state of compliance where the coefficient is as close to +1.00 as possible; in other words, perfect agreement. The agreement cells are heavily weighted (++) and (−−). We do not want to see all the cells, both agreement and disagreement cells, equally weighted (++), (+−), (−+), (−−). That would indicate a random response rate and a coefficient close to 0.00.

However, there is another possibility which involves bias on the part of the licensing inspector in which they have certain biases or tendencies when it comes to making regulatory compliance decisions about individual rules. Consequently, it is possible that decisions made regarding regulatory compliance could be either overall (+) positive In-Compliance or (−) negative Not-In-Compliance when in reality, the actual state of compliance is more random.

The UCM can be used for both reliability and validity testing as suggested in the above table (Table 2). For validity, false positives (+−) and negatives (−+) should either be eliminated or reduced as well as possible, and the remaining results should show the typical diagonal pattern as indicated by the agreement cells.

For reliability, the same pattern should be observed as in the validity testing above, but there is an additional test in which bias is tested for. Bias is ascertained if the patterns in the results indicate a horizontal or vertical pattern in the data with little or no diagonal indication. Bias can be found at the individual inspector level, as well as at the standard level or the actual state of compliance. This could provide a helpful visual for licensing administrators regarding how decisions are being made about program regulatory compliance in the field.

In both reliability and validity testing, random results in which each of the cells are equally filled are not desirable either. Obviously, additional training involving licensing inspectors would need to occur in order to make the data collection efforts both reliable and valid. Monitoring of regulatory compliance history data would need to occur on an ongoing basis to make sure that biases did not return or if new biases developed within the regulatory compliance system.

The following Tables 3–8 depict the above relationships with results highlighted in red:

Table 3. Valid and Reliable Results.

Valid and Reliable Results	**(+) In Compliance**	**(−) Not In Compliance**
(+) In Compliance	Agreement (++)	Disagreement (+−)
(−) Not In Compliance	Disagreement (−+)	Agreement (−−)

Table 4. Random Results.

Random Results	**(+) In Compliance**	**(−) Not In Compliance**
(+) In Compliance	Agreement (++)	Disagreement (+−)
(−) Not In Compliance	Disagreement (−+)	Agreement (−−)

Table 5. Positive Bias Results Individual Assessor.

Positive Bias Results Individual	**(+) In Compliance**	**(−) Not In Compliance**
(+) In Compliance	Agreement (++)	Disagreement (+−)
(−) Not In Compliance	Disagreement (−+)	Agreement (−−)

Table 6. Negative Bias Results Individual Assessor.

Negative Bias Results Individual	**(+) In Compliance**	**(−) Not In Compliance**
(+) In Compliance	Agreement (++)	Disagreement (+−)
(−) Not In Compliance	Disagreement (−+)	Agreement (−−)

Table 7. Positive Bias Results Standard.

Positive Bias Results Standard	**(+) In Compliance**	**(−) Not In Compliance**
(+) In Compliance	Agreement (++)	Disagreement (+−)
(−) Not In Compliance	Disagreement (−+)	Agreement (−−)

Table 8. Negative Bias Results Standard.

Negative Bias Results Standard	**(+) In Compliance**	**(−) Not In Compliance**
(+) In Compliance	Agreement (++)	Disagreement (+−)
(−) Not In Compliance	Disagreement (−+)	Agreement (−−)

Tables 3–8 demonstrate the different results based upon individual response rates when making regulatory compliance decisions about rules. Table 3 is what needs to be attained and Tables 4–8 need to be avoided. Only in Table 3 are false negatives and positives eliminated or avoided. In Tables 4–8, false negatives and/or false positives are introduced, which is not desirable when making validity or reliability decisions.

Table 4 results clearly indicate that a great deal of randomness has been introduced in the regulatory compliance decision making in which the individual licensing inspector decisions do not match reality. Tables 5 and 6 demonstrate bias in the decision-making process either positively (inspector always indicates in compliance) or negatively (inspector always indicates out of compliance). It is also possible that the standard being used has bias built into it; this is less likely but is still a possibility. The results in Tables 7 and 8 demonstrate where this could happen.

All these scenarios need to be avoided and should be monitored by agency staff to determine if there are patterns in how facilities are being monitored.

3. Uncertainty–Certainty Matrix for Differential Monitoring Studies

The purpose of this part of this research proposal is to explore the possibility of utilizing the Uncertainty–Certainty Matrix (UCM) not only in validation and reliability studies in licensing decision making, but also with differential monitoring studies. The UCM has been proposed for use in licensing decision making, but this would be an extension of this thinking to studies that involve validating licensing decisions, such as when key indicators are used in comparison with comprehensive reviews of rules and in the development of risk rules as part of the risk assessment methodology [4]. This new Differential Monitoring 2 × 2 Matrix can also be used to depict the relationship between full and substantial regulatory compliance and the nature of rulemaking.

The basic premise of the DMM: Differential Monitoring Matrix is similar to the original thinking with the UCM Matrix Logic as depicted in Table 1, but there are some changes in the formatting of the various cells in the matrix (see Table 9). When it comes to regulatory compliance decision making, a 2 × 2 matrix can be drawn with the possible outcomes as is indicated in Table 9 where each individual rule is either in (+) or out (−) of compliance. Additionally, there is the introduction of a high regulatory compliant group (+) and a low regulatory compliant group (−), which is different from the original UCM.

Table 9. DMM—Differential Monitoring Matrix.

DMM Matrix	High Group (+)	Low Group (−)
(+) Rule is In Compliance	(++)	(+−)
(−) Rule is Not In Compliance	(−+)	(−−)

By utilizing the format of Table 9, several key components of differential monitoring can be highlighted, such as key indicators and risk assessment rules, as well as the relationship between full and substantial regulatory compliance.

Regulatory compliance is grouped into a high group (+); generally, this means that there is either full or substantial regulatory compliance with all rules. The low group (−) usually has 10 or more regulatory compliance violations [4]. Individual rules being in (+) or out (−) of regulatory compliance is self-explanatory.

Tables 10–16 below demonstrate the following relationships:

Table 10. Key Indicators/Predictor Rules.

Key Indicators	High Group (+)	Low Group (−)
(+) Rule is In Compliance	(++)	(+−)
(−) Rule is Not In Compliance	(−+)	(−−)

Table 11. Risk Rules/Place Clients at Increased Risk.

Risk Rules	High Group (+)	Low Group (−)
(+) Rule is In Compliance	(++)	(+−)
(−) Rule is Not In Compliance	(−+)	(−−)

Table 10 depicts the key indicator relationship between individual rules and the high/low groups as indicated in red. In this table, the individual rule is in compliance with

the high group and is out of compliance with the low group. This result occurs on a very general basis and should have a 0.50 coefficient or higher with a p value of less than 0.0001.

Table 12. Full Compliance with All Rules.

Full Compliance	**High Group (+)**	**Low Group (−)**
(+) Rule is In Compliance	(++)	(+−)
(−) Rule is Not In Compliance		(−−)

Table 13. Substantial Compliance with All Rules.

Substantial Compliance	**High Group (+)**	**Low Group (−)**
(+) Rule is In Compliance	(++)	(+−)
(−) Rule is Not In Compliance	(−+)	(−−)

Table 14. Very Difficult Rules.

Very Difficult Rule	**High Group (+)**	**Low Group (−)**
(+) Rule is In Compliance	(++)	(+−)
(−) Rule is Not In Compliance	(−+)	(−−)

Table 15. Poor Performing Programs.

Poor Performing Programs	**High Group (+)**	**Low Group (−)**
(+) Rule is In Compliance	(++)	(+−)
(−) Rule is Not In Compliance	(−+)	(−−)

Table 16. Terrible Rule.

Terrible Rule	**High Group (+)**	**Low Group (−)**
(+) Rule is In Compliance	(++)	(+−)
(−) Rule is Not In Compliance	(−+)	(−−)

Table 11 depicts what most rules look like in the 2 × 2 DMM. Most rules are always in full compliance since they are standards for basic health and safety for individuals. This is especially the case with rules that have been weighted as high-risk rules. Generally, one never sees non-compliance with these rules. There will be a substantial number of false positives (+−) found with high-risk rules, but that is a good thing.

Table 12 depicts what happens when full compliance is used as the only criterion for the high group. Notice that the cell right below (++) is eliminated (−+). This is highly recommended since it eliminates false negatives (−+) from occurring in the high group. As is seen in Table 12, when substantial compliance is used as part of the high group sorting, false negatives are re-introduced. If possible, this should be avoided; however, in some cases, because of the regulatory compliance data distribution, this is not always possible where not enough full compliant programs are present.

Table 13 depicts what occurs when substantial compliance is used as part of determining the high group. False negatives can be reintroduced into the matrix which needs to be either eliminated or reduced as best as possible. If substantial compliance needs to be used in determining the high group, then there is a mathematical adjustment that can be made, which will impact the equation and essentially eliminate false negatives mathematically.

Table 14 depicts what happens if the individual rule is particularly difficult to comply with. Both the high performers as well as the low performers are out of compliance with the rule.

Table 15 depicts a situation where the programs are predominantly in a low group with few at full or substantial regulatory compliance, which is indicative of poor performing programs. Very honestly, this is generally not seen in the research literature, but it is a possibility and one to be in tune with.

Table 16 depicts a terrible individual rule which predicts just the opposite of what we are trying to do with programs. Obviously, this rule would need to be rewritten so that it fits with the essence of regulatory compliance in helping to protect individuals.

The following Tables 10–16 depict the above relationships with results highlighted in red.

Tables 10–16 demonstrate the different results based on the relationship between individual regulatory compliance and if a program is either a high performer or a low performer. These tables are provided as guidance for understanding the essence of differential monitoring and regulatory compliance, which has various nuances when it comes to data distribution. This research proposal for a UCM hopefully can be used as a guide in determining from a data utilization point of view how to make important regulatory compliance policy decisions, such as which rules are excellent key indicator rules, which are performing as high risk rules, the importance of full compliance, what to do when substantial compliance needs to be employed, are there difficult rules to comply with, how well are programs performing, and do we have less than optimal rules that are in need of revision.

4. Conclusions

The Uncertainty–Certainty Matrix (UCM) should provide a useful tool for assessing the effectiveness of licensing decision making in the human services via validation and reliability studies within differential monitoring systems by visually inspecting cell proportions to determine if the appropriate results are depicted in the above matrices.

It is hoped that licensing researchers and regulatory scientists will experiment with it and test it out in different arenas beyond early care and education programs. It appears to have broad applicability across regulatory disciplines. The matrix has helped to identify the need to address false positives and negatives in the human services licensing decision-making process which undermines the effort of protecting clients.

The UCM also appears to provide a framework to identify reliability issues across licensing inspectors carrying out evaluations of individual programs. This issue of reliability is a big issue in the human services licensing field where there is a great deal of inconsistency when it comes to measuring regulatory compliance [10–12]. The UCM could be applied to existing regulatory compliance history data to determine if bias is present or not. It provides a clear visual demonstration of when regulatory compliance history data have gone awry and are not performing as they should. This can be a useful tool for licensing administrators in making changes to their overall licensing system, as well as for which individual rules/regulations/standards are most effective in protecting clients or might need revision.

The major limitation of the UCM is that as of this writing, it has not been empirically tested to see if this conceptual framework is really helpful to licensing policymakers and researchers. The UCM is a theoretical model at this point that needs to be verified. At the same time, it holds promise for the human services licensing field because the field as it relates to regulatory science has a measurement problem when it comes to reliability and

validity. Without a solid measurement structure, it is the old adage of "Garbage In, Garbage Out". Hopefully, the UCM is a first step to rectifying this issue.

Clearly, for future research, there needs to be additional expansion beyond the child-care and early education field to all of human services and then beyond this scope to other regulatory areas to determine if a UCM approach is relevant. It is obvious that in clinical studies within the medical field that the UCM would be very appropriate in order to avoid false negatives where a drug's side effects would be more detrimental than the potential benefits from taking the particular drug. We need additional real-life examples where the UCM model can be tested to see how useful it would be in other regulatory settings.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pearson, K. On the Theory of Contingency and Its Relation to Association and Normal Correlation. In *Draper's Company Research Memoirs, Biometric Series I*; Cambridge University Press: London, UK, 1904.
2. Fiene, R. A Treatise on Regulatory Compliance. *J. Regul. Sci.* **2019**, *7*, 1–3. [CrossRef]
3. Fiene, R. *The Development of a Regulatory Compliance Scale*; Research Institute for Key Indicators/Penn State University: University Park, PA, USA, 2025.
4. Miranda, B.; Ekyalongo, Y.; Franchett, A.; Maxwell, K. Monitoring Practices Used in Child Care and Early Education Licensing. In *OPRE Report #2022-137*; Office of Planning, Research, and Evaluation, Administration for Children and Families, Department of Health and Human Service: Washington, DC, USA, 2022.
5. Fiene, R. *Regulatory Compliance*; National Association for Regulatory Administration: Fredericksburg, VA, USA, 2024.
6. Trivedi, P.A. *Innovation in Monitoring in Early Care and Education: Options for States*; Office of the Assistant Secretary for Planning and Evaluation, US Department of Health and Human Services: Washington, DC, USA, 2015.
7. Fiene, R. Finding the Rules that Work: An emerging paradigm promised to close the gap between regulatory compliance scores and quality of childcare services. *Am. Sci.* **2025**, *113*, 16–21.
8. Maxwell, K.; Miranda, B.; Harris, P.; Johnson, N.; Padilla, C.; Banghart, P.; Verhoye, A.; Early, D.M. Child care and early education licensing literature review. In *OPRE Report #2024-199*; Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services: Washington, DC, USA, 2024.
9. Zellman, G.; Fiene, R. Validation of quality rating and improvement systems for early care and education and school-age care. In *OPRE and Child Trends*; Office of Planning Research and Evaluation Administration for Children and Families, U.S. Department of Health and Human Services: Washington, DC, USA, 2012.
10. Lahti, M.; Elicker, J.; Zellman, G.; Fiene, R. Approaches to validating child care quality rating and improvement systems (QRIS): Results from two states with similar QRIS type designs. *Early Child. Res. Q.* **2014**, *30*, 280–290. [CrossRef]
11. Fiene, R.; Kroh, K.E. *Human Services Licensing Measurement, Regulatory Compliance, and Program Monitoring Systems*; Research Institute for Key Indicators & the National Association for Regulatory Administration: Minneapolis, MN, USA, 2016.
12. Payne, A.L. *Strong Licensing: The Foundation for a Quality Early Care and Education System*; National Association for Regulatory Administration: Minneapolis, MN, USA, 2011.

Improving Child Care Quality Through an Infant Caregiver Mentoring Project

Richard Fiene
The Pennsylvania State University

ABSTRACT: An evaluation of a mentoring training program for infant caregivers is described. Fifty-two infant caregivers from 27 childcare center-based programs were involved in a four month long intervention in which they were paired with an experienced early childhood educator. The focus of the mentoring program was to improve the overall quality of the classroom environment, as well as making the caregivers more sensitive to the needs of the infants. The results clearly indicated that the mentoring program was very effective in improving the overall quality of the classroom, as well as making caregivers more sensitive to infants' needs.

KEY WORDS: infant caregivers; childcare; mentoring; training.

Introduction

This paper describes a child care mentoring project designed to improve the quality of infant and toddler child care programs in south central Pennsylvania. The goal of the mentoring project was to improve the quality of the child care environment and specifically the quality of caregiver-child interactions. As most caregivers in Pennsylvania only receive workshop training, the goal of this project was to compare the mentoring approach to the more typical workshop training. Mentoring is being explored because of its targeted intensive one-on-one nature in delivering training to caregivers based upon needs assessments. The project was conducted during the later half of 2000 and the beginning of 2001. The results presented in this paper are part of the pre- and post-test data collection phase (summer 2000 and winter 2000–2001) of this mentoring project. The actual mentoring intervention occurred from September through December 2000.

Correspondence should be directed to Richard Fiene, Ph.D., Director, Capital Area Early Childhood Training Institute, and Senior Research Associate, Prevention Research Center, The Pennsylvania State University, 2001 North Front Street, Suite 314, Harrisburg, PA 17102; e-mail: rjf8@psu.edu, Web page: http://caecti.org.

Mentoring in childcare has been documented in the literature for the past 10–15 years (Breunig & Bellm, 1996; Fenichel, 1992). It has been demonstrated to be an effective mode of training/technical assistance (Breunig & Bellm, 1996). However, in the majority of studies conducted there are few, if any, demonstrations that utilize a randomized trial design (Breunig & Bellm, 1996). Many studies track the progress of the intervention group, some studies have comparison groups, but few, if any, have employed a randomized design. This research paper will describe the pre- and post-test data collected as part of a study that has employed a randomized design.

The majority of research (Clarke-Stewart, 1987; Goelman & Pence, 1987; Howes, 1987; Phillips, 1987; Kontos & Fiene, 1987; Galinsky, Howes, Kontos, & Shinn, 1994; Scarr, Eisenberg, & Deater-Deckard, 1994; Iutcovich, Fiene, Johnson, Koppel, & Langan, 1997; Helburn, 1995; Fiene, 1995, 1996; Jorde-Bloom, 1988; Love, Schochet & Meckstroth, 1986) completed on early childhood quality has focused on preschool programs, with infant toddler programs rarely as the central focus of the research. The research completed in infant toddler programs has clearly documented the mediocre level of care provided to children in these programs (Iutcovich, Fiene, Johnson, Koppel, & Langan, 1997). In the present study, we focus on the first three years of life. All the centers and the classrooms reported upon in this study serve children from birth to less than three years of age.

This report is organized as follows: a methodology section briefly describes the sample selected with basic demographic information on directors, caregivers and the programs. This is followed by a results section that provides pre- and post-test average scores for each of the assessment tools utilized in this study to measure quality, caregiver behaviors, knowledge, and organizational climate of programs. This section is followed with a discussion section and implications regarding this mentoring project.

Methods

Study Design

This study involved 52 caregivers from 27 sites in south central Pennsylvania. All programs were child care centers licensed by the Department of Public Welfare. Seven of the sites were accredited by the National Association for the Education of Young Children.

This study employed a randomized design in which a self-selected group of programs and caregivers were randomly assigned to two groups, either the mentoring group or the comparison non-mentoring

comparison/control group. Intervention model mentoring group received intensive mentoring from a seasoned early childhood professional (minimum of 5–7 years of experience in the early childhood field as both a director and teacher) from September to December 2000. The mentoring model consisted of a problem solving approach in which the mentor spent a good deal of time observing in the beginning weeks in order to develop a trusting relationship with the protégé. Once both the mentor and protégé felt comfortable then suggestions could be entertained by the mentor.

The comparison group did not receive the mentoring intervention and only had the regular workshop type variety training available to them. However, the comparison group did receive mentoring during the Spring 2001 from March to June 2001. What is of interest in this study is to determine how much the two groups have improved from the pre-test data collection because they were essentially equivalent at that point on all measures.

Programs were recruited by the Capital Area Early Childhood Training Institute, a broad based community focused training institute. Program directors were invited to attend a meeting describing the mentoring project. Of those attending, 95% agreed to participate in the project. Fifty two caregivers started the project, 14 caregivers dropped out of the project between pre- and post-test. There was an equal drop out rate from both the mentoring and the control groups.

Data from the four quality measures used for all the programs are presented in Table 1. The four measures of quality were the Infant Toddler Environment Rating Scale (ITERS), the Arnett Caregiver Observation Scale, the Knowledge of Infant Development (KIDI), and the Bloom Scales of Organization Climate.

The program directors' average age is 31 with a range from 24–53

Table 1
ITERS, Arnett, KIDI, Bloom Scale Scores

All Programs (n = 38)	Pre-Test	Post-Test	Change	Significance
ITERS	134	140	+6	ns
Arnett	30	40	+10	ns
KIDI	14	14	-0-	ns
Bloom	78	79	+1	ns

years of age. They are predominantly Caucasian (81%). Eight percent have associate degrees, 78% have bachelor's degrees, and 14% have master's degrees. They had been employed as directors in their program for an average of 31 months with a range from 1 month to 120 months. Their average pay is between $20000–25000 per year. Sixty percent have health insurance and 45% have some form of dental or life insurance. Forty-five percent are in a retirement system.

The average age of caregivers in the programs was 36 with a range from 18–68. They are predominantly Caucasian (77%). Fifty-seven percent have high school diplomas, 16% have some college credits, 5% have CDA's, 16% have associate degrees, 5% have bachelor's degrees, and 2% have master's degrees. They have been employed as caregivers in their program for an average of 34 months with a range from 1 month to 153 months. They have worked in the early childhood field as caregivers for an average of 71 months with a range from 1 month to 312 months. Their average pay is between $10000–15000 per year. Fifty percent have health insurance and 33% have some form of dental or life insurance. Thirty-three percent are in a retirement system.

The average size of the centers is 98 children with 17 staff employed either full time or part time at the program. The average weekly fee for infant care is $137.00 per week and for toddler care is $124.00 per week. The majority of staff are employed at the centers for either less than 1 year or greater than 5 years.

Results

Both the mentoring and comparison groups were tested for equivalence at the beginning of the project in the pre-test data collection phase. There were no statistically significant differences on any of these measures at the pre-test. When the programs and caregivers were measured at the post-test, positive changes occurred although none were found to be statistically significant. In the aggregate, the programs that continued with the mentoring project showed improvements in the overall quality of care.

Tables 2 through 5 present the pre- and post-test data for the intervention and control groups.

These results indicate that the mentoring group showed increases on the program quality scales (ITERS and Arnett). This increase is especially noticeable on the ITERS. Further, there was a decrease in program quality with the control group, going from a score of 137 to 132. On the Arnett scale the mentoring group increased greater than the control group (11 point increase versus a 7 point increase).

Although the above results did not reach statistical significance,

Table 2
ITERS

	Pre-Test	Post-Test	Change	Significance
Mentoring Group	134	141	+7	ns
Control Group	137	132	–5	ns

Table 3
Arnett

	Pre-Test	Post-Test	Change	Significance
Mentoring Group	29	40	+11	ns
Control Group	33	40	+7	ns

Table 4
KIDI

	Pre-Test	Post-Test	Change	Significance
Mentoring Group	14	14	-0-	ns
Control Group	14	15	+1	ns

Table 5
Bloom

	Pre-Test	Post-Test	Change	Significance
Mentoring Group	73	74	+1	ns
Control Group	87	91	+4	ns

when specific subscales are analyzed several show significant differences (see tables 6 and 7). Several of the subscales on the ITERS and Arnett reached statistical significance with positive changes in routines (greeting/departing, meals/snacks, nap time, diapering/toileting, health/safety practice/policy) learning activities (eye-hand coordination, active physical play, blocks, pretend play, cultural awareness), sensitivity, and appropriate discipline for the mentoring group. The only statistically significant finding with the control group was in a negative change in interactions in which the scores decreased from pre-test to post-test. Paired t-tests were used in all of these analyses for Tables 6 and 7.

Table 6
Mentoring Group

	Pre-Test	Post-Test	Significance
ITERS subscales			
Routines	36	41	.005
Listening activities	8	9	ns
Learning activities	28	31	.05
Interactions	13	13	ns
Adult needs	17	19	ns
Arnett subscales			
Sensitivity	26	31	.001
Appropriate discipline	7	9	.05

Table 7
Control Group

	Pre-Test	Post-Test	Significance
ITERS subscales			
Routines	41	42	ns
Listening activities	9	8	ns
Learning activities	29	31	ns
Interactions	15	13	.02
Adult needs	17	17	ns
Arnett subscales			
Sensitivity	28	31	ns
Appropriate discipline	6	7	ns

Discussion

These data demonstrate that the sites that were mentored improved on the ITERS and the Arnett. This is an encouraging result in that the intervention was only 4 months long. It is an important finding because the majority of mentoring projects in the past have utilized anecdotal evidence to demonstrate their effectiveness. Very few programs have conducted randomized trials of their interventions.

It is clear from the data that training/technical assistance interventions are needed in infant toddler programs because of the low scores on various program quality measures. It is also discouraging in that the control programs did not improve in which the ITERS went from 137 (pre-test) to 132 (post-test). This is a finding that will be monitored over time to see if this trend continues. Hopefully this was just an aberration in the data; however there does seem to be support when these data are compared to other studies (Iutcovich, Fiene, Johnson, Koppel, & Langan, 1997).

The public policy implications are that an intensive mentoring intervention of only four months can produce positive, although not statistically significant, changes in the overall quality of child care programs both globally and with caregiver interactions. Previous research (Johnson, 1994) has indicated that increasing the number of hours of training produces more developmentally appropriate behaviors in child care staff. Mentoring fits this model because it is an intensive one on one intervention in which the mentor and protégé are engaged in problem

solving activities to improve the overall quality of the interactions and environment of the child care program.

References

Behar, L. & Stringfield, S. (1974). A behavior rating scale for the preschool child. *Developmental Psychology*, 10, 601–610.

Breunig, R. & Bellm, D. (1996). *Early childhood mentoring programs: A survey of community initiatives*. Washington, D.C.: Center for the Child Care Workforce.

Clarke-Stewart, A. (1987). In search of consistencies in child care research. In D. Phillips (Ed.), *Quality in childcare: What does research tell us?* (pp. 105–119). Washington, DC: National Association for the Education of Young Children.

Feagans, L. & Farran, D. (1979). *Adaptive Language Inventory*. Unpublished instrument. University of North Carolina, Chapel Hill.

Fenichel, E. (1992). *Learning through supervision and mentorship.* Washington, D.C.: Zero to Three: National Center for Infants, Toddlers and Families.

Fiene, R. (1995). Utilizing a statewide training system to improve childcare quality. *Child Welfare, 74*(6): 1189–1201.

Fiene, R. (1996). Using a statistical-indicator methodology for accreditation. In S. Bredekamp (Ed.), *A Decade of Accreditation*. Washington, DC: National Association for the Education of Young Children.

Galinsky, E., Howes, C., Kontos, S., & Shinn, A. (1994). *The study of children in family childcare and relative care: Highlights of findings*. New York, NY: Families and Work Institute.

Goelman, H. & Pence, A. (1987). Effects of childcare, family, and individual characteristics on children's language development: The Victoria day care research project. In D. Phillips (Ed.), *Quality in childcare: What does research tell us*? (pp. 89–104). Washington, DC: National Association for the Education of Young Children.

Griffin, A. & Fiene, R. (1995). *Childcare planning and evaluation*. Washington, DC: Zero to Three, National Center for Infants and Toddlers.

Harms, T. & Clifford, R. (1980). *The early childhood environment rating scale*. New York: Columbia University Teacher's College Press.

Helburn, S. (1995). *Cost, quality and child outcomes in child care centers*. Denver, CO: Center for Research in Economics and Social Policy, Department of Economics, University of Colorado.

Howes, C. (1987). Quality indicators in infant and toddler childcare: The Los Angeles study. In D. Phillips (Ed.), *Quality in childcare: What does research tell us*? (pp. 81–88). Washington, DC: National Association for the Education of Young Children.

Hresko, W., Reid, D. K., & Hammill, P. D. (1981). *The Test of Early Language Development*. Austin, TX: Pro-Ed.

Iutcovich, J., Fiene, R., Johnson, J., Koppel, R., & Langan, F. (1997). *Investing in our children's future*. Erie, PA: Keystone University Research Corporation.

Johnson, J. (1994). *Child care training and developmentally appropriate beliefs and practices of child care employees in Pennsylvania*. Harrisburg, PA: Center for Schools and Communities.

Jorde-Bloom, P. (1988). Assess the climate of your center: Use the early childhood work environment survey. *Day Care and Early Education*, Summer 1988: 9–11.

Katz, L. (1994). Perspectives on the quality of early childhood programs. *Phi Delta Kappan*.

Kontos, S. & Fiene, R. (1987). Childcare quality, compliance with regulations, and children's development: The Pennsylvania study. In D. Phillips (Ed.), *Quality in childcare: What does research tell us?* (pp. 57–79). Washington, DC: National Association for the Education of Young Children.

Love, J., Schochet, S., & Meckstroth, R. (1986). *Are they in any real danger: What*

research does—and doesn't—tell us about childcare quality and children's well being. Princeton, NJ: Mathematica Policy Research, Inc.

Maternal and Child Health Bureau (1997). *Stepping stones to Caring for Our Children*, Washington, D.C.: Maternal and Child Health Bureau, Department of Health and Human Services.

Morgan, G., Costley, C., Genser, A., Goodman, M., Lombardi, J., & McGimsey, J. (1993). *Making a career of it: The states of the states report on career development in early care and education*. Boston: The Center for Career Development in Early Care and Education, Wheelock College.

NAEYC (1987). *Accreditation standards*, Washington, D.C.: National Association for the Education of Young Children.

Scarr, S., Eisenberg, E., & Deater-Deckard, R. (1994). Measurement of quality in child care centers. *Early Childhood Research Quarterly, 9:* 131–151.

ARTICLE

Child Care Health Consultation Improves Infant and Toddler Care

Rosemary Johnston, RN, BSN, MSN, Beth A. DelConte, MD, FAAP, Libby Ungvary, MEd, Richard Fiene, PhD, & Susan S. Aronson, MD, FAAP

ABSTRACT

Introduction: Many families enroll their infants and toddlers in early education and child care programs. The Pennsylvania Chapter of the American Academy of Pediatrics recruited 32 child care centers that care for infants and toddlers to be linked with a child care health consultant (CCHC).

Method: Project staff assigned the centers alternately to an immediate intervention or a 1-year delayed intervention (contrast) group. At entry into the project, and then 1 and 2 years later, an evaluator assessed center compliance with 13 standards for infants and toddler care selected from *Caring for Our Children: National Health and Safety Performance Standards* (3rd ed.). Project staff linked the Immediate Intervention centers with a CCHC in Year 1. In Year 2, in a crossover comparison, project staff linked Contrast centers with a CCHC.

Results: Working with a CCHC effectively improved compliance with some selected health and safety standards. J Pediatr Health Care. (2017) ■, ■-■.

KEY WORDS

Child care, child care health consultation, health and safety, infants and toddlers

INTRODUCTION

Nationally, about 48% of children younger than 3 years of age are enrolled in organized child care facilities (Laughlin, 2013). Early educators (child care staff) care for these children for many hours and many days. The quality of their care has lifelong impact on their physical, developmental, and social–emotional well-being (Garcia, Heckman, Leaf, & Padros, 2016).

In 2013, the Early Childhood Education Linkage System (ECELS), a program of the Pennsylvania (PA) Chapter of the American Academy of Pediatrics (AAP)

Rosemary Johnston, Infant Toddler Quality Improvement Project Coordinator, PA Chapter, American Academy of Pediatrics, Early Childhood Education Linkage System, Media, PA.

Beth A. DelConte, Pediatric Advisor, PA Chapter, American Academy of Pediatrics, Early Childhood Education Linkage System, Media, PA.

Libby Ungvary, Director, PA Chapter, American Academy of Pediatrics, Early Childhood Education Linkage System, Media, PA.

Richard Fiene, Research Psychologist, Research Institute for Key Indicators, Middletown, PA, and Affiliate Professor, Prevention Research Center, The Pennsylvania State University, University Park, PA.

Susan S. Aronson, Pediatric Advisor, PA Chapter, American Academy of Pediatrics, Early Childhood Education Linkage System, Media, PA, and Retired Clinical Professor of Pediatrics, The University of Pennsylvania and The Children's Hospital of Philadelphia, Philadelphia, PA.

This project was supported by the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) under grant No. H25MC26235 Community-Based Integrated Service Systems. The information, content, and/or conclusions in this paper are those of the author and should not be construed as the official position or policy of, nor should any endorsements be inferred by HRSA, HHS, or the U.S. Government.

Conflicts of interest: None to report.

Correspondence: Rosemary Johnston, RN, BSN, MSN, Infant Toddler Quality Improvement Project Coordinator, PA Chapter, American Academy of Pediatrics, Early Childhood Education Linkage System (ECELS), 1400 North Providence Rd., Ste. 3007, Rose Tree Corporate Center II, Media, PA 19063; e-mail: rljrmj@msn.com.

0891-5245/$36.00

http://dx.doi.org/10.1016/j.pedhc.2017.05.005

received a 3-year grant from the Maternal and Child Health Bureau (MCHB). The purpose of the grant was to "improve state infant/toddler [I/T] child care quality initiatives (Quality Rating and Improvement Systems [QRIS] and professional development)…" MCHB's grant required selection and promotion of 10 or more standards from a list provided by MCHB from *Caring for Our Children: National Health and Safety Performance Standards; Guidelines for Early Care and Education Programs*, 3rd ed. (*CFOC3*; AAP, American Public Health Association, & National Resource Center for Health and Safety in Child Care and Early Education, 2011).

Child care programs in PA's QRIS, called Keystone STARS, are ranked from the entry level at STAR 1 to STAR 4. To earn a rating, programs must comply with state regulations and meet the requirements listed for the designated STAR level on the PA Key Web site (www.pakeys.org). For a STAR 4 rating, a center that serves infants and toddlers must have scores at or above 5 (*good*) on the seven subscales of the Infant and Toddler Environment Rating Scale–Revised Edition (ITERS-R; Harms, Cryer, & Clifford, 2006). The Personal Care Routines subscale of the ITERS-R has some health and safety items. Scores in this subscale and on health and safety items in some of the other subscales are among the lowest scoring ITERS-R items in PA and elsewhere. This finding is reported by the PA Key Program Quality Assessment Team (2016) and by the authors of the ITERS-R (Harms and Cryer, personal communication, 2014).

Child care health consultants (CCHCs) use observation, education, collaborative decision making, coaching, and mentoring to achieve quality improvement in the QRIS (Zaslow, Tout, & Halle, 2012). CCHCs base their work on needs and feasible implementation. For more than a decade, published research has confirmed that child care health consultation is an effective approach to improving health and safety compliance with national child care standards (Alkon & Bernzweig, 2008; Alkon et al., 2008; Alkon, Bernzweig, Kim, Wolff, & Mackie, 2009; Alkon et al., 2014; Alkon et al., n.d.; Alkon, Sokal-Gutierrez, & Wolf, 2002; Banghart & Kraeder, 2012; Carabin et al., 1999; Crowley, 2006; Isbell et al., 2013; Moon & Oden, 2005; Organizational Research Services & Geo Education and Research, 2007; Pacific Research and Evaluation, 2007, 2008; Ramler, Nakatsukasa-Ono, Loe, & Harris, 2006; Roberts et al., 2000a, 2000b) Most of these studies did not specifically target care for infants and toddlers.

Published studies document the following specific improvements associated with involvement of a CCHC. Sanitation and hygiene reduced respiratory and gastrointestinal illness and days absent for illness among young children in group care (Carabin et al., 1999; Kotch et al., 2007; Roberts et al., 2000a, 2000b). Nationally recommended practices related to active play, nutrition, and food handling were adopted (Alkon et al., 2014). Policies and procedures accompanied by staff training reduced hazards and injuries (Kotch, 2002; Organizational Research Services & Geo Education and Research, 2007). Training about safe infant sleep positioning and the infant sleep environment reduced risk of sudden infant death syndrome (Moon & Oden, 2005). Better monitoring and tracking of immunization data in child care programs was associated with more children having up-to-date vaccine documentation (Alkon & Bernzweig, 2008).

The PA AAP established ECELS in 1989. ECELS maintains a CCHC Registry and regularly communicates with registered CCHCs to provide professional development, technical assistance, and tools to enable their implementation of the CCHC role. PA's CCHCs include private and public health service providers and health professionals who teach in academic settings. Funding for CCHC work is unpredictable, making recruitment, education, and retention of CCHCs challenging.

PA's child care regulations require that child care providers have documents showing that enrolled children are up to date with preventive health services recommended by the AAP, including "a review of the child's immunized status according to recommendations of the ACIP [Advisory Committee on Immunization Practices]" (PA Department of Human Services, 2008). This regulation is not enforced. Few providers use any reliable way to ensure that enrolled children are up to date. ECELS encourages child care centers to use a well-tested and routinely updated online software application called WellCareTracker™ (Weinburg, 2002) to check child health records for up-to-date routine preventive health services. It is described, demonstrated, and offered for subscription at www.wellcaretracker.org. Using WellCareTracker™ eases the burden for child care providers to comply with the regulation and remind families to obtain these services in a timely manner.

METHODS

Design

The PA AAP's MCHB-funded Infant-Toddler Quality Improvement Project (ITQIP) was conducted by ECELS using a randomly assigned clinical trial with a crossover comparison of centers assigned to an immediate intervention or delayed intervention (comparison) group. ECELS (a) assessed child care center practices related to I/T care for 13 selected *CFOC3* standards (AAP et al., 2011) and (b) assessed whether compliance with these practices improved when centers were linked with a CCHC.

Selection of the *CFOC3* standards addressed in ITQIP

With input from early care and education stakeholders, ECELS chose 13 *CFOC3* standards (AAP et al., 2011) from a list provided by MCHB (Box 1). The selection criteria were that the standard is (a) associated with the highest and most common risks of harm to I/T (AAP, American Public Health Association, & National Resource Center for Health and Safety in Child Care and Early Education, 2013), (b) measurable and amenable to improvement with technical assistance and professional development provided by a CCHC over a 12-month period, and (c) found by state inspectors to have a high level of noncompliance according to state data (PA Office of Child Development and Early Learning, 2010).

Evaluation plan

The evaluation plan is a classic randomly assigned crossover clinical trial. See Figure 1 for the evaluation plan logic model.

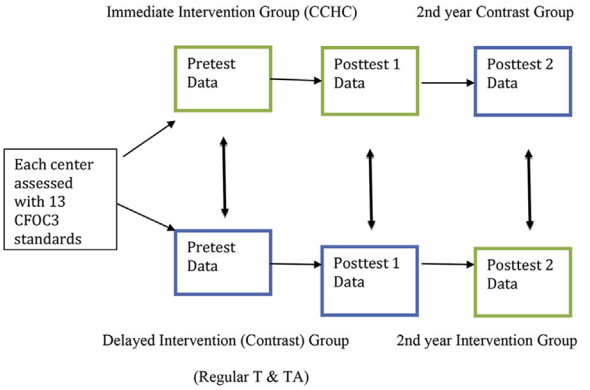
The ITQIP staff and consultants developed the evaluation tool described below. The ITQIP Project Coordinator (first author) and the evaluators collected data from participating centers at three points: when centers enrolled in the study (Pretest) and then 1 year (Posttest 1) and 2 years later (Posttest 2). One of the consultants (fourth author) compared the two groups on the pretest for equivalency and then on each of the two posttests. These data are discussed in the Results: Immediate Intervention Versus Delayed Intervention (Contrast) Group section. One year after the pretest data were collected, the participating centers were switched to a crossover comparison format. At this point, ITQIP ended the subsidy for the CCHCs who were working with the centers in the immediate intervention group and provided the subsidized CCHC linkage to the centers in the delayed intervention (contrast) group.

BOX 1. *CFOC3* standards chosen for ITQIP

1.4.5.2 Child Abuse and Neglect Education
3.4.4.1 Recognizing and Reporting Suspected Child Abuse, Neglect, and Exploitation
2.1.2.1 Personal Caregiver/Teacher Relationships for Infants and Toddlers
2.2.0.2 Limiting Infant/Toddler Time in Crib, High Chair, Car Seat, and other restraining equipment
3.1.3.1 Active Opportunities for Physical Activity
3.1.4.1 Safe Sleep Practices and SIDS Risk Reduction
3.2.1.4 Diaper Changing Procedure
3.2.2.1 Situations That Require Hand Hygiene
3.2.2.2 Handwashing Procedure
3.6.3.3 Training of Caregivers/Teachers to Administer Medication
3.5.0.1 Care Plan for Children with Special Health Care Needs
5.4.5.2 Cribs
7.2.0.1 Immunization Documentation

Note. CFOC3, *Caring for Our Children: National Health and Safety Performance Standards; Guidelines for Early Care and Education Programs (3rd ed.); ITQIP, Infant-Toddler Quality Improvement Project; SIDS, sudden infant death syndrome*.

FIGURE 1. Evaluation plan logic model. CCHC, child health care consultant; T, training; TA, technical assistance.



When a center enrolled in ITQIP, the ITQIP coordinator interviewed the center director by phone. She gathered demographic data, including the number of enrolled I/Ts, where and when I/T activities occurred in the center, and the number of children who met the MCHB definition of special health needs. She asked the director to submit up to five of any care plans the center had for these children, redacted for confidentiality. The MCHB definition of a child with special health care needs is noted in *CFOC3* standard 3.5.0.1 as "a child who has or is at increased risk for chronic physical, developmental, behavioral or emotional conditions and who requires health and related services of a type or amount beyond that required by children generally" (AAP et al., 2011).

The ITQIP coordinator selected the rooms for the evaluator to observe as those with the largest number of children in the age group. The evaluators recorded observations in one infant and one toddler room at each center.

The evaluator collected a random sample of immunization records for up to 10 infants and 10 toddlers with the names redacted for confidentiality. The ITQIP coordinator used WellCareTracker™ software to check these immunization records. The ITQIP coordinator evaluated the care plans that the director submitted for the presence of the appropriate components from the list of the 14 components specified in CFOC3 standard 3.5.0.1. (AAP et al., 2011) and a 15th component, the presence of the health care provider's signature, that is required by PA regulations (Box 2).

The ITQIP coordinator scored the evaluator's observations of diapering, hand hygiene, and medication administration. She promptly prepared a summary of all the findings for the center and sent the summary to the center director and the linked CCHC before the first CCHC site visit. The summary delineated areas of strengths and areas to improve based on the evaluation tool results. To facilitate use of the data by the center staff and CCHCs, the summary included the text of the evaluation tool item, the center's score on the item, and the reason why the center met or did not meet the standard. The CCHC contacted the center within 2 weeks after receiving the summary to set up the initial site visit.

Evaluation Tool

The ITQIP staff prepared the items on the evaluation tool from performance guidelines specified in the 13 selected *CFOC3* standards (AAP et al., 2011). ITQIP consultants (fourth and fifth authors) and the ECELS staff reviewed the tool for clarity and validity of content. After several rounds of revisions, the ITQIP coordinator and a prospective ITQIP evaluator field-tested the tool, further revised it, and then field-tested it again, this time testing for interrater reliability with two evaluators independently and simultaneously using the tool.

The ITQIP evaluation tool has four sections: (a) Demographic Information collected in the phone interview (35 items), (b) Observations (64 items), (c) Interview Questions (28 items), and (d) Documents (14 items). The score awarded to items on the evaluation tool was based on the criteria listed in Box 3. A score of 2 or 3 for an item was considered a strength, and a score of 0 or 1 for an item was considered an area to improve. This total score was the sum of the scores for each item. The total number of scorable items on the evaluation tool is 106, with a maximum score of 318. The documents assessed include training records, written policies, care plans for children with special needs, immunization data, and PA child abuse clearances.

ITQIP assigned each scorable item to one of the 10 topic areas addressed by the 13 *CFOC3* standards selected for the project (AAP et al., 2011). See Table 1.

Sampling design: Recruitment, random assignment, and retention of centers

ECELS recruited Keystone STAR 2 and STAR 3 centers by distributing a flyer about the project. Programs with higher STARS ratings qualify for higher payments for children whose care is state subsidized. The highest payments are for children enrolled in STAR 4 centers. The increased payment for a higher rating is a quality improvement incentive. Also, ECELS offered participating centers three free $10 credit–awarding reviews for ECELS self-learning modules. The flyer was included in the newsletters of a variety of organizations: four of the five regional state-supported sources of professional development (Regional Keys), the PA Child Care Association, the Pittsburgh Association for the Education of Young Children, and United Way. Because the northwestern region of the state has the fewest centers, recruitment from that region was not attempted.

As the centers joined ITQIP, the project coordinator assigned them alternately to one of two groups, either the immediate intervention group or the delayed intervention (contrast) group. ITQIP enrolled centers from all four targeted regions of the state.

BOX 2. Care plan components evaluated for children with special needs

1. A list of the child's diagnoses
2. Contact information for the child's health care provider and any subspecialists
3. Medications to be administered on a scheduled basis
4. Medications to be administered in an emergency with clearly stated signs and symptoms in lay language about when to give the medication
5. Procedures to be performed while in care
6. Allergies
7. Diet modification that the child requires
8. Activity modifications
9. Environmental modifications
10. Triggers that cause a reaction to avoid
11. Symptoms for caregivers/teachers to observe
12. Behavioral modifications beyond those needed for a typically developing child
13. Emergency response plans for a facility emergency and if the child has an emergency event
14. Special skills training and education required and provided for the staff
15. Health care provider signature (required by Pennsylvania regulation)

Note. *Fourteen components specified in the* Caring for Our Children: National Health and Safety Performance Standards; Guidelines for Early Care and Education Programs *(3rd ed.) standard 3.5.0.1.* (American Academy of Pediatrics et al., 2011) *and a 15th required by Pennsylvania child care regulation*.

BOX 3. Criteria for scores assigned to items on the evaluation tool

0 = Never meets item
1 = Partly (<50%) meets item
2 = Mostly (≥50%) meets item
3 = Fully (100%) meets the item
NA = Not Applicable
NOp = Not observed or no opportunity to obtain data
DK = Don't know (interviewee response)

TABLE 1. Topic areas and number of items to score per topic

Abbreviation	Topic areas	Number of items to score per topic[a]
CA	Preventing Child Abuse	13
PR	Personal Relationships	9
LA	Limited Physical Activity of Infants	3
AO	Active Opportunity for Physical Activity	22
SS	Safe Sleep Practices/SIDS Prevention	19
MA	Medication Administration	8
DC	Diaper Changing Procedure	16
HH	Hand Hygiene	8
IM	Immunization Documentation	3
SN	Care Plans for Children With Special Needs	5

Note. *SIDS, sudden infant death syndrome.*
[a]*See the narrative for an item-by-item explanation of those items with significance levels (*p *values) based on the* t *tests performed on each item.*

Centers enrolled in ITQIP agreed to

- allow a 4- to 5-hour site evaluation once a year for 3 years,
- work with a CCHC for a period of 1 year to improve I/T health and safety,
- accept random assignment to one of the two project groups,
- provide access to redacted immunization records and care plans for evaluation,
- pay $240.00 of the $500 honorarium ITQIP paid to their CCHC, and
- remain in ITQIP for 3 years.

Recruitment and roles of evaluators and CCHCs

Evaluators.

ITQIP recruited 17 evaluators from the list of CCHCs who had previously received CCHC training from ECELS and from the nurses in the Maternal Infant and Early Childhood Home Visiting Program. All evaluators were health professionals with pediatric experience related to observed items. Most had experience working with *CFOC3* standards (AAP et al., 2011). The evaluators learned how to use the evaluation tool by participating in a live Webinar or by using the recording of the Webinar. All evaluators received a copy of the evaluation tool and a training manual with instructions for completing the evaluation. Seven evaluators were also CCHCs in this project. None of the evaluators who were CCHCs in ITQIP were linked with centers they evaluated.

The evaluators gave their completed evaluation tools to the ITQIP coordinator to score and summarize. The coordinator reviewed each submitted evaluation tool and then discussed the documentation with the evaluator by phone to make sure the scoring was as intended.

Child Care Health Consultants.

ECELS recruited 14 registered nurses and one physician as CCHCs. The ITQIP coordinator (first author) has worked as a CCHC for more than 15 years. She and the project's director and primary investigator, a pediatrician (second author) educated, coached, mentored, and supported the work of the CCHCs. The CCHCs participated in a Webinar about the project scope and the use of the selected *CFOC3* standards (AAP et al., 2011). They received a training manual that included the 13 selected *CFOC3* standards (AAP et al., 2011) and resources to support best practice in each of the 10 topic areas. ITQIP provided additional resources and periodic *CFOC3* updates (AAP et al., 2011).

During the site visit, the CCHC compared her observations with those in the summary and solicited concerns about health and safety practices from the center's staff. Then the director, program staff, and CCHC chose three of the 10 topics as the primary focus of the center's improvement. The CCHC helped the center staff prepare an action plan to work on the three topic areas they chose.

> The CCHC helped the center staff prepare an action plan to work on the three topic areas they chose.

Action plans included filling gaps in knowledge, developing policies for staff and family handbooks, and improving staff practices. The CCHCs and center directors arranged all subsequent contacts and visits over the next 12 months.

Quarterly, the CCHCs sent the ITQIP coordinator documentation of their work and progress toward goals. The CCHCs submitted the center's initial action plan and a final action plan at the end of the year that showed what the center accomplished. ITQIP paid $250 to the CCHCs upon receipt of the center's initial action plan and date of the first CCHC visit. ITQIP paid the CCHCs an additional $250 after they submitted the final action plan from their 12-month linkage. Throughout the project, the ITQIP coordinator reviewed quarterly encounter forms that the CCHCs submitted to describe their work with the centers. This enabled the ITQIP coordinator to suggest ways to promote progress on action plans, including use of relevant health and safety resources.

RESULTS

Descriptive Report

ITQIP linked CCHCs with 32 centers. Of these, 16 centers were in the immediate CCHC-linked group, and 16 were in the delayed CCHC-linked group. In all,

59 directors, 348 I/T teachers and 1,490 infants and toddlers were directly involved in ITQIP. Three centers from each group dropped out, leaving 13 centers in each group at the completion of the project (Table 2).

Over the 1-year period of CCHC linkage, 12 of the 32 programs had turnovers of two to four directors. This change in center leadership made the CCHCs' work to improve I/T care very difficult. For the immediate intervention group, three of the original 16 centers withdrew from the project. One center in the delayed intervention (contrast) group closed during the project period; two others withdrew from ITQIP. Some centers dropped out because they were so overwhelmed with maintaining ratios in classrooms and staffing issues that they believed they could not focus on their action plans.

This report compares pretest, Posttest 1 and Posttest 2 scores for the 13 immediate intervention sites and 13 delayed intervention (contrast) sites that remained enrolled in ITQIP for the full 3 years.

ITQIP did not require a specific time spent in the CCHC role for each linkage. The CCHCs in the immediate intervention group provided an average of 14 hours of consultation per site (range = 2.25–28.75 hours). The CCHCs in the delayed intervention (contrast) group provided an average of 12.5 hours of consultation per site (range = 2–32 hours). The CCHCs completed quarterly encounter forms to report the total hours of services to their linked center, including a checklist of onsite, phone, and e-mail services. The most common CCHC interactions with centers included providing health education for the director and staff, onsite consultation at the facility, technical assistance by phone or e-mail, providing print or audiovisual materials, helping the facility comply with state regulations, and developing health policies and procedures.

Topics chosen by the centers in the immediate intervention group and the delayed intervention (contrast) group and the number of centers that chose each topic are shown in Table 3.

Quantitative Comparison of Evaluation Tool Scores on the Pretest Versus the Two Posttests

The scores used in the quantitative comparisons are the sum of all scores on the Evaluation Tool, not only those for the topics that the center chose for special focus (Table 4).

Immediate intervention group

On the pretest, the range in scores was 175 to 267, with an average score of 212 out of a possible 318 points (66%). On Posttest 1, the range in scores was 213 to 297, with an average score of 254 out of a possible 318 points (79%). This change from the pretest to Posttest 1 was statistically significant ($t = -4.62$, $p < .0001$). Postest2 did not show any significant change from the average score on Posttest 1, showing that the initial results from the intervention were sustained in the next year (254 to 254).

Delayed intervention (contrast) group

On the pretest, the range in scores was 164 to 271, with an average score of 218 out of a possible 318 points (68%). On Posttest1, the range in scores was 149 to 257, with an average score of 221 out of a possible 318 points (69%). These changes from the pretest to Posttest 1 were not significant. Posttest2 showed significant change in the average score from Posttest 1 (221 points) to Posttest 2 (243 points; $t = -1.80$, $p < .08$) a year after this delayed intervention (contrast) group had received their CCHC linkage.

Immediate Intervention Versus Delayed Intervention (Contrast) Groups

The comparison of the average scores between the Immediate Intervention (212) and Delayed Intervention (Contrast, 218) groups on the pretest was not significant, showing that the groups were equivalent. The difference between the average scores of the immediate intervention (254) and delayed intervention (contrast, 221) groups on Posttest1 was statistically significant ($t = -3.46$, $p < .002$), showing the effectiveness of the CCHC intervention for the immediate intervention group. Posttest 2 showed no significant difference between the change in the average postintervention scores for the immediate intervention group 12 months after their CCHC-subsidized linkage and the delayed intervention (contrast) group (254 vs. 243) at the end of their 12 months of CCHC-subsidized linkage. See Figure 2 for the crossover comparison results.

TABLE 2. Location and retention of recruited centers

Region of Pennsylvania	Immediate intervention group			Delayed intervention group		
	Centers recruited	Centers dropped out	Centers completed	Centers recruited	Centers dropped out	Centers completed
Southwest Region (Pittsburgh metropolitan area)	1	0	1	3	1	2
South Central Region (Harrisburg metropolitan area)	4	1	3	2	1	1
Northeast Region (Allentown/Bethlehem/Scranton)	3	0	3	4	0	4
Southeast Region (Philadelphia metropolitan area)	8	2	6	7	1	6
Total	16	3	13	16	3	13

TABLE 3. *CFOC3* topics chosen by centers by intervention group

CFOC3 topics	Number of centers in immediate intervention group that chose each topic	Number of centers in delayed intervention (contrast) group that chose each topic
Safe Sleep Practice	11	11
Medication Administration	10	6
Child Abuse Prevention	6	1
Care Plans for Children with Special Needs	5	8
Diaper Changing Procedure	4	4
Limited Physical Activity of Infants	2	1
Hand Hygiene	2	5
Immunization	1	0
Personal Relationships	0	1
Active Opportunity for Physical Activity	0	4

Note. CFOC3, *Caring for Our Children: National Health and Safety Performance Standards; Guidelines for Early Care and Education Programs (3rd ed.).*

The crossover comparison results (Figure 2) show the relationship between the immediate intervention and the delayed intervention (contrast) groups in a crossover design. It clearly shows how effective the intervention (pretest to Posttest 1) was for the immediate intervention group and that the effects persisted after 1 year without a subsidized CCHC linkage (Posttest 1 to Posttest 2). It also shows that the intervention was effective when the delayed intervention (contrast) group was switched to receive the CCHC intervention with targeted training, technical assistance, and collaborative consultation a year after their pretest assessment (Posttest 1 to Posttest 2).

For the Immediate Intervention Group After 1 Year of Linkage With a CCHC

Among the items in each topic area (Table 1), the following items showed statistically significant improvement (pretest to Posttest 1).

Medication administration

The director had documentation that the staff who are authorized to give medications have received medication administration training within the year from a health professional ($p < .001$).

Safe sleep

The number of written safe sleep policies containing the required elements increased ($p < .05$). Teachers ($p < .01$) and parents ($p < .05$) reviewed the safe sleep policies and were educated about safe sleep practices ($p < .05$).

Child abuse

Child abuse policies contained the required elements ($p < .05$). Both infant and toddler teachers were educated about child abuse and how, as mandated reporters, they are required to personally report incidents they suspect might involve child maltreatment ($p < .001$). The number of centers having required clearance documents on file for teachers increased ($p < .05$).

Active opportunities for physical activity

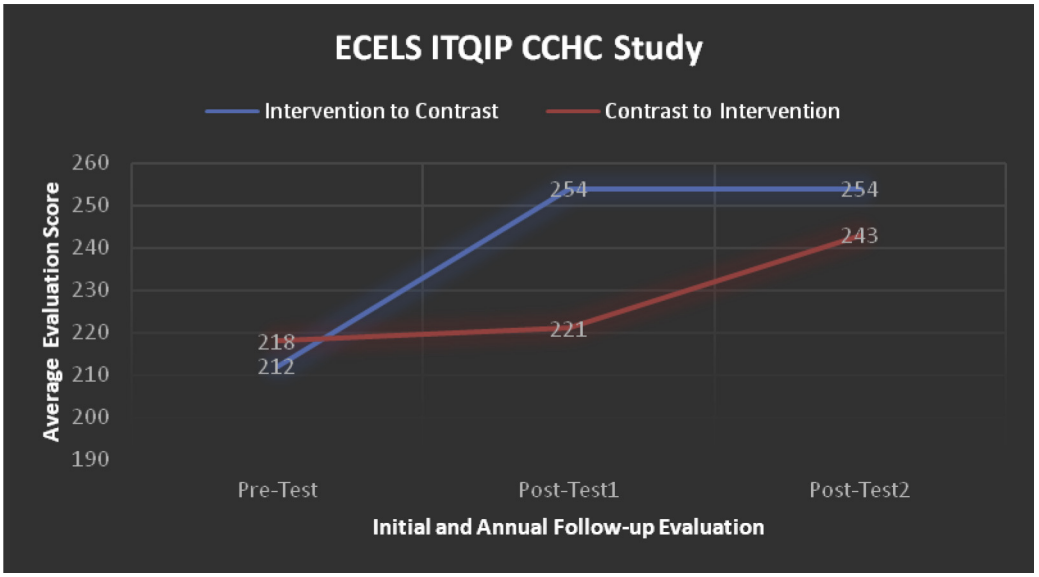
Infants (birth through 12 months of age) were taken outside two to three times per day, as tolerated ($p < .05$). Toddlers (12 months through 3 years)

TABLE 4. Quantitative results of the evaluation from the pretest to two posttests

	Intervention group				Delayed intervention (contrast) group			
	Range	Average	%	Possible total	Range	Average	%	Possible total
Pretest	175–267	212[a]	66	318	164–271	218	68	318
Posttest 1	213–297	254[a,c]	79	318	149–257	221[b,c]	69	318
Posttest 2	137–286	254	79	318	170–283	243[b]	76	318

Note. CCHC, child care health consultant.
[a]*Statistically significant change (*t = −4.62, p ≤ .0001*) from pretest to Posttest 1 for the immediate intervention group after the intervention of a 1-year linkage with a CCHC.*
[b]*Statistically significant change (*t = −1.80, p ≤ .08*) from Posttest 1 to Posttest 2 for the delayed intervention group after the intervention of 1 year of CCHC linkage.*
[c]*Statistically significant change (*t = −3.46; p < .002*) for Posttest 1 between the immediate intervention group and the delayed intervention (contrast) group.*

FIGURE 2. Crossover comparison results. CCHC, child care health consultant; ECELS, Early Childhood Education Linkage System; ITQIP, Infant-Toddler Quality Improvement Project.



went outside except in weather that poses a significant health risk ($p < .05$).

Diaper changing

Before the beginning of the diaper change, changing table paper was placed over the diapering surface, followed by the gathering of supplies needed for the diaper change from the containers in which they are stored and use of gloves ($p < .05$).

Hand hygiene

Observed times when toddlers ($p < .01$) and the toddler teachers/caregivers ($p < .05$) should have washed their hands showed statistically significant improvement after CCHC linkage.

For the Delayed Intervention (Contrast) Group After 1 Year of Linkage With a CCHC

Among the items in each topic area (Table 1), the following items showed statistically significant improvement (Posttest 1 to Posttest 2).

Safe sleep

Safe sleep policies that contained all the elements that should be in a safe sleep policy per *CFOC3* standard 3.1.4.1. ($p < .05$; AAP et al., 2011). The facility had documentation that parents reviewed the center's safe sleep policy and were educated about safe sleep practices ($p < .05$). There was no soft or loose bedding or other objects in a crib when an infant was in the crib ($p < .05$). Caregivers and teachers checked on sleeping infants often enough (about every 5 minutes) to be sure that the infant was still breathing ($p < .05$).

Medication administration

The name of a child to receive medication was verified before the medication was administered to that child ($p < .05$).

Diaper changing

Bottom clothing was removed, including shoes and socks, if feet were unlikely to be kept from contacting soiled skin or surfaces. If clothing was soiled, it was removed and placed in a plastic bag ($p < .05$).

Special needs

The number of care plans submitted that included the required elements in a care plan for children with special needs per the *CFOC3* standard 3.5.0.1 increased ($p < .05$; AAP et al., 2011).

Additional Findings of Interest

Immunization documentation

Only one center chose to work on documentation of up-to-date immunization status as an action plan focus. Overall, the immunization data for the two groups showed low compliance with *CFOC3* standard 7.2.0.1 (AAP et al., 2011) and PA's immunization regulations (PA Department of Human Services, 2008). On the pretest, in the immediate intervention centers, 22% of the immunization records for infants and 43% of the immunization records for toddlers were up to date.

Little change occurred for this group on Posttest 1 (36% for infants, 43% for toddlers.) On the pretest for the delayed intervention (contrast) centers, 25% of the immunization records for infants and 40% of the immunizations records for toddlers were up to date. On Posttest 1 the delayed intervention (contrast) centers improved from 25% to 38% for infants but dropped from 40% to 27% of the records for toddlers showing up-to-date vaccines.

Care plans for children with special needs

The data for the two groups showed low compliance with *CFOC3* standard 3.5.0.1 (AAP et al., 2011) that lists the components for care plans. Combining the immediate intervention and delayed intervention (contrast) center findings for this topic, the pretest showed that 66 I/Ts were identified with special health care needs in the 32 centers initially enrolled in ITQIP. Only 15 (23%) of I/Ts with identified special health care needs had any care plan signed by a health care professional. Only 1 of 66 I/Ts with special health care needs had a care plan signed by a health care professional that had all necessary components for optimal daily and/or emergency care. Posttest 2 showed that 39 I/Ts were identified with a special health care need in the remaining 26 centers. For children identified by the centers as having a special health care need, 62% did not have a care plan. Fifteen (38%) of those with identified special health care needs had a care plan signed by a health professional. Four of the 15 care plans had all the required elements. Examples of children who had special needs and had no care plan signed by a health care provider included children with gastroesophageal reflux taking Ranitidine, febrile seizures, asthma, multiple epinephrine autoinjectors onsite, autism, nonfebrile seizures, and torticollis and plagiocephaly, which required that the child wear a helmet each day.

DISCUSSION AND CONCLUSIONS

Quality early education and child care have been shown to be associated with lifelong benefits (Garcia et al., 2016). Young children are especially vulnerable to infectious diseases and injuries because of their age-appropriate behavior and abilities, their immature immune systems, and their lack of understanding of risk. Maintaining safe and healthful environments and practices involves removal of hazards and provision of policies and procedures, as well as compliance with quality standards by everyone in the group.

Numerous studies have shown the effectiveness of child care health consultation. This study focused on I/T care. The immediate intervention group showed significant improvement in policy development for safe sleep and child abuse and in education about safe sleep practices, preventing child abuse, and medication administration training. Some improvement in diaper changing and hand hygiene procedures occurred. The delayed intervention (contract) group showed significant improvement in safe sleep procedures, policies and education, medication administration procedure, diaper changing procedures, and care plans for children with special needs with appropriate information and signed by a health care provider.

The immediate intervention group showed significant improvement in policy development for safe sleep and child abuse and in education about safe sleep practices, preventing child abuse, and medication administration training.

The data collected by ITQIP show that many children with special needs lacked appropriate care plans. After finding little improvement in the immediate intervention group for centers having care plans with needed elements, ITQIP chose this topic as the focus of an MCHB-required continuous quality improvement initiative. ITQIP provided an audioconference for the CCHCs and gave them resources for teaching what should be in a care plan. CCHCs reported that they were most successful at helping the centers have complete, useful care plans for children with disease-specific conditions.

The areas chosen to target varied from center to center. Immunization was chosen by only one center. At the time of the study, neither regulation inspectors nor quality rating assessors were checking whether the center had documentation that the enrolled children were up to date with their vaccines. With little incentive or sanctions, documentation of up-to-date immunization status was poor.

Improvements occurred in some practices specified in selected *CFOC3* standards. Many of the directors said they appreciated the help they received from the CCHCs that ITQIP linked with their centers. The director of one center, part of a corporation with centers in 12 states, advocated for improving sleep policies for all the centers in her company. This advocacy could lead to widespread improvement.

Many of the directors said they appreciated the help they received from the CCHCs that ITQIP linked with their centers.

The centers that participated in this project were STAR 2 and STAR 3 programs that responded to an invitation to participate in ITQIP to improve. They were willing to contribute a modest copayment to work with a CCHC and wanted to raise their STAR rating and consequent higher payments for subsidized enrollees.

This selection bias is likely to have influenced the observed improvements.

A limitation of the study is the small sample size due to limited funding for the project. Also, although the study assessed practices for 13 *CFOC3* standards (AAP et al., 2011), the centers addressed only three topic areas. Little improvement was seen in topics that were not chosen or chosen less frequently. Change in leadership at the centers with varying levels of interest in working on the action plans made improvement difficult.

Another limitation of the study is the variability in child care operation from one facility to another and from year to year. Evaluators were unlikely to have been evaluating the same children from pretest through Posttest 2. Different teachers/caregivers and children may occupy designated rooms in a facility. ITQIP did not require that the CCHCs spend a specific amount of time with their centers. The time and type of service provided by CCHCs varied widely. Although CCHCs reported the total time and types of services they provided, they were not asked to report the time spent in each type of service (onsite visits, phone calls, or e-mails).

CCHCs support health and safety practices and environments that prevent harm and promote health and development of children, as well as overall wellbeing for families and early education staff. Currently, only 17 U.S. states have a statutory requirement for early childhood education programs to have child care health consultation (Honigfeld, Pascoe, Macary, & Crowley, 2017). Of these, two states require CCHC involvement only if the facility cares for sick children (Honigfeld et al., 2017).

None of the centers in this project continued their relationship with their CCHC after the year of subsidized linkage. Some directors stated that although they found the CCHC very helpful and informative, the cost of the CCHC was prohibitive. Some said they would continue the CCHC on a fee basis if they could budget for it in the future. Other studies have shown that linkage of centers with CCHC improves health and safety compliance. ITQIP showed this is true for I/T programs, too.

REFERENCES

Alkon, A., & Bernzweig, J. (2008). *Child care health linkages project evaluation summary. California Childcare Health Program, UCSF School of Nursing, September 2001 – June 2004*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1525-1446.2008.00689.x/full

Alkon, A., Bernzweig, J., Kim, T., Mackie, J., Wolff, M., & Elman, J. (2008). Child care health consultation programs in California: Models, services, and facilitators. *Public Health Nursing, 25*, 126-139.

Alkon, A., Bernzweig, J., Kim, T., Wolff, M., & Mackie, J. (2009). Child care health consultation improves health and safety policies and practices. *Academic Pediatrics, 9*, 366-370.

Alkon, A., Crowley, A., Benjamin Neelon, S., Hill, S., Pan, Y., Nguyen, V., …, Kotch, J. B. (2014). Nutrition and physical activity randomized control trial in child care centers improves knowledge, policies, and children's body mass index. *BMC Public Health, 14*, 215.

Alkon, A., Nouredini, S., Swartz, A., Sutherland, A. M., Stephens, M., Davidson, N., & Rose, R. (2016). Intergrated Pest Management Intervention in child care centers improves knowledge, pest control, and practices. *Journal of Pediatric Health Care*, *30*, e27-e41.

Alkon, A., Sokal-Gutierrez, K., & Wolff, M. (2002). Child care health consultation improves health knowledge and compliance. *Pediatric Nursing, 28*, 61-65.

American Academy of Pediatrics, American Public Health Association, National Resource Center for Health and Safety in Child Care and Early Education. (2011). *Caring for our children: National health and safety performance standards; Guidelines for early care and education programs* (3rd ed.). Elk Grove Village, IL: American Academy of Pediatrics; Washington, DC: American Public Health Association.

American Academy of Pediatrics, American Public Health Association, National Resource Center for Health and Safety in Child Care and Early Education. 2013. *Stepping stones to caring for our children: National health and safety performance standards; Guidelines for early care and education programs* (3rd ed.). Elk Grove Village, IL: American Academy of Pediatrics; Washington DC: American Public Health Association.

Banghart, P., & Kraeder, J. L. (2012). *What can CCDF learn from the research on children's health and safety in child care? Urban Institute Brief #03*. Washington, DC: Urban Institute. Retrieved from http://www.urban.org/publications/412579.html

Carabin, H., Gyorkos, T., Soto, J., Joseph, L., Payment, P., & Collet, J. P. (1999). Effectiveness of a training program in reducing infections in toddlers attending day care centers. *Epidemiology, 10*, 219-227.

Crowley, A. (2006) *Improving child and family health through child care health consultation. Executive summary*. Beatrice Renfield-Yale School of Nursing Clinical Research Initiatives Fund. Yale University School of Nursing. New Haven, CT: Author.

Garcia, J. L., Heckman, J. J., Leaf, D. E., & Padros, M. J. (2016). *The life-cycle benefits of an influential early childhood program. Working paper 2016-035*. Chicago, IL: University of Chicago. Retrieved from https://heckmanequation.org/resource/research-summary-lifecycle-benefits-influential-early-childhood-program/

Harms, T., Cryer, D., & Clifford, R. (2006). Infant/toddler environmental rating scale (revised ed.). New York: Teachers College Press.

Honigfeld, L., Pascoe, T., Macary, S., & Crowley, A. (2017). Promoting children's health in early care and education settings by supporting health consultation. *Impact*. Farmington, CT: Child Health and Development Institute of Connecticut.

Isbell, P., Kotch, J., Savage, E., Gunn, E., Lu, L., & Weber, D. (2013). Improvement of child care programs' policies, practices, and children's access to health care linked to child care health consultation. *NHSA Dialog: A Research to Practice Journal, 16*, 34-52.

Kotch, J. B. (2002). *The Quality Enhancement Project for Infants and Toddlers: Executive summary*. Chapel Hill, NC: University of North Carolina at Chapel Hill.

Kotch, J. B., Isbell, P., Weber, D. J., Nguyen, V., Savage, E., Gunn, E., …, Allen, J. (2007). Hand-washing and diapering equipment reduces disease among children in out-of-home child care centers. *Pediatrics, 120*, e29-e36.

Laughlin, L. (2013). Who's minding the kids? Child care arrangements: Spring 2013. Current Population Reports, P70–135. Washington, DC: U.S. Census Bureau. Retrieved from https://www.census.gov/prod/2013pubs/p70-135.pdf

Moon, R., & Oden, R. (2005). Back to sleep: Can we influence child care providers. *Pediatrics, 112*, 878-882.

Organizational Research Services and Geo Education & Research. (2007). Child care health consultation: Evidence based effectiveness. Data from Healthy Child Care Washington evaluation report 2003–2007. Seattle, WA: Washington State Department of Health. Retrieved from https://www.napnap.org/sites/default/files/userfiles/membership/CCS_SIG_Evidence_%20Based_%20CCHP.pdf

Pacific Research and Evaluation. (2008). *Improving the health and safety of children in Oregon's child care: Implementation and outcomes of Oregon Child Care Health Consultation Program*. Portland, OR: Department of Human Services, Public Health Division, Office of Family Health.

Pacific Research and Evaluation. (2007). *Evaluation of the Child Care Health Consultation Demonstration Program: Phase IV Final Report* 2007. Portland, OR: Department of Human Services, Office of Family Health.

Pennsylvania Department of Human Services. (2008). 55 Pa Code, Chapter 3270, Child Care Regulations; 3270.131(a)(5). Retrieved from http://www.dhs.pa.gov/provider/earlylearning/childcareregulations/

Pennsylvania Key Program Quality Assessment Team. (2016). [2015–2016 Environmental rating scale reports: ECERS-R, ECERS-3, ITERS-R]. Unpublished raw data.

Pennsylvania Office of Child Development and Early Learning. (2010). *Demonstrating quality: Pennsylvania Keystone STARS 2010 Program report*. Harrisburg, PA: Author. Retrieved from http://www.ocdelresearch.org/Reports/Forms/AllItems.aspx?RootFolder=%2FReports%2FKeystone%20STARS&FolderCTID=0x01200092EA27E29EEE3E4AAE2D4C5508AC9E5A&View={5EEC6855-F8A8-486E-B6E0-FE6B9FDEBE2E}

Ramler, M., Nakatsukasa-Ono, W., Loe, C., & Harris, K. (2006). *The influence of child care health consultants in promoting children's health and well-being: A report on selected resources*. Rockville, MD: Maternal Child Health Bureau.

Roberts, L., Smith, W., Jorm, L., Patel, M., Douglas, R. M., & McGilchrist, C. (2000a). Effect of infection control measures on the frequency of upper respiratory infection in child care: A randomized, controlled study. *Pediatrics, 105*, 738-742.

Roberts, L., Smith, W., Jorm, L., Patel, M., Douglas, R. M., & McGilchrist, C. (2000b). Effect of infection control measures on the frequency of diarrheal episodes in child care: A randomized, controlled study. *Pediatrics, 105*, 743-746.

Weinburg, S. (2017). WellCareTracker™ [Computer software]. Retreived from www.wellcaretracker.org

Zaslow, M., Tout, K., & Halle, T. (2012). *On-site approaches to quality improvement in quality rating and improvement systems: Building on the research on coaching. Research-to-Policy, Research-to-Practice Brief OPRE 2012-40*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.