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.

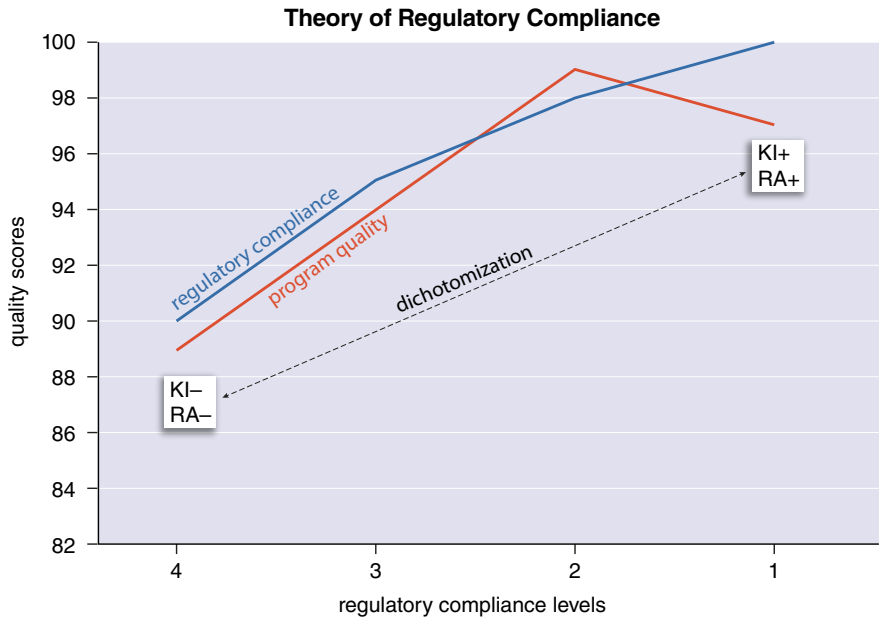
But first, some background: As the U.S. Department of Health, Education, and Welfare took over running the show for all U.S. early care and education programs in the 1970s, *uniform program monitoring* had become the rule. Uniform monitoring derived from the philosophical assumption that fuller regulatory 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

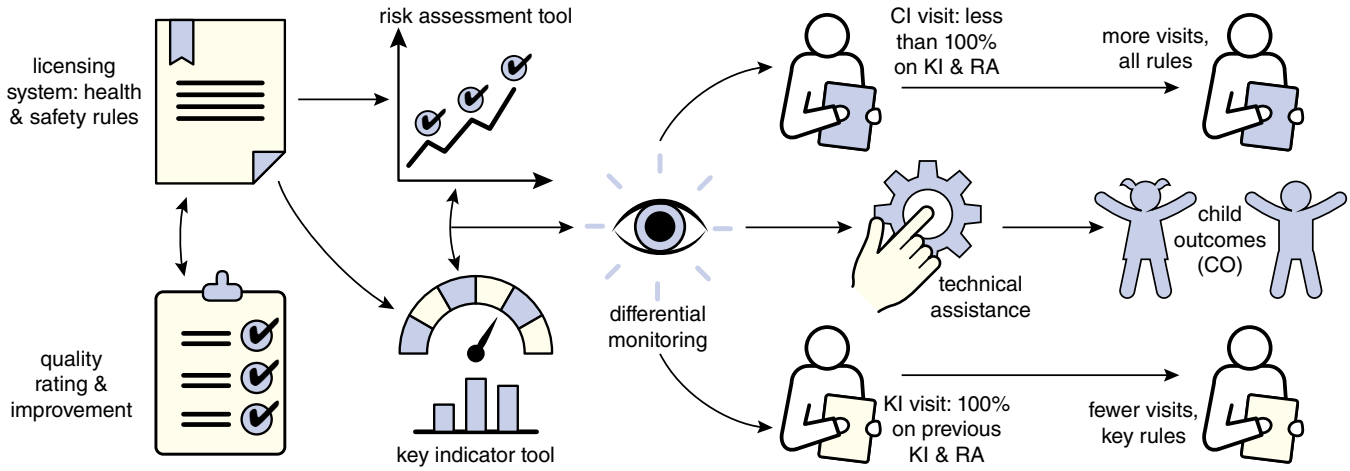
This table compares different approaches to measuring compliance: A licensing-focused approach in which programs are classified as either compliant or noncompliant based on rules violation counts, with no middle ground (*columns 1 and 2*), and a more nuanced ordinal approach using a Likert scale. This 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

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.

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.

Published in *Scholarly Community Encyclopedia*, April, 2025, pps 1-20.

Development of a Regulatory Compliance Scale: A New Licensing Metric

Richard Fiene PhD

Research Institute for Key Indicators

The Penn State Edna Bennett Pierce Prevention Research Center

September 2025

Development of a Regulatory Compliance Scale: A New Licensing Metric

Richard Fiene PhD

Penn State's Edna Bennett Pierce Prevention Research Center

September 2025

Abstract

Introduces a new licensing metric called the Regulatory Compliance Scale which moves licensing measurement from a nominal to an ordinal measurement strategy. This new scale provides an enhanced metric for regulatory compliance and comparing licensing reviews with other more process quality measures available in the child care and early education field. This is a major change within human services licensing. This article presents the possibilities of this new scale and its limitations as well as the logic behind the scale and the studies that have been done to date to support the use of the Regulatory Compliance Scale.

Keywords Licensing measurement; regulatory compliance scale; theory of regulatory compliance; structural quality; process quality; nominal measurement; ordinal measurement.

Introduction

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 (Fiene, 2025c). 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 (NAEYC, 2025). 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 (Fiene, 2019, 2022, 2025a).

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 being suggested 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)(Harms etal, 2023; Fiene, 2025a). 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.

Regulatory Compliance Scale: The Logic Model

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, 2019, 2025a). 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.

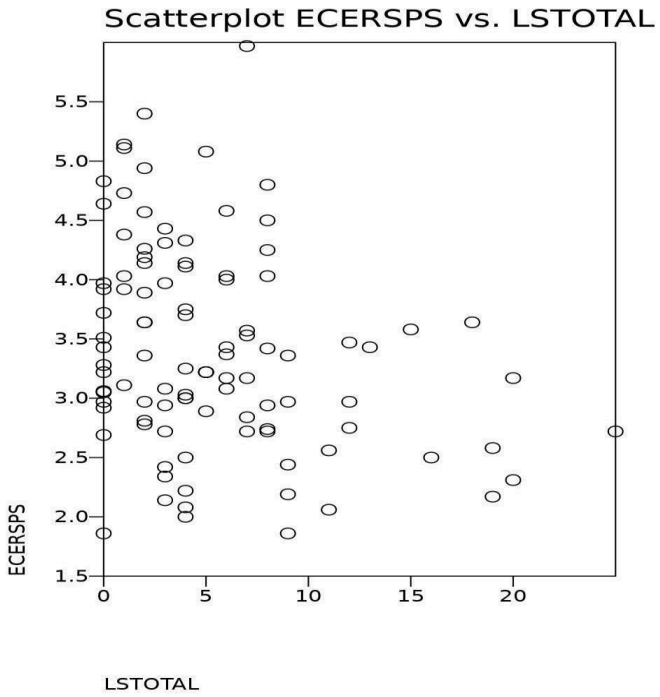
Table 1: Regulatory Compliance Scale (RCS) Compared to Risk and Violations

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. These studies were completed in various jurisdictions in the USA and in Canada (NARA, 2023).

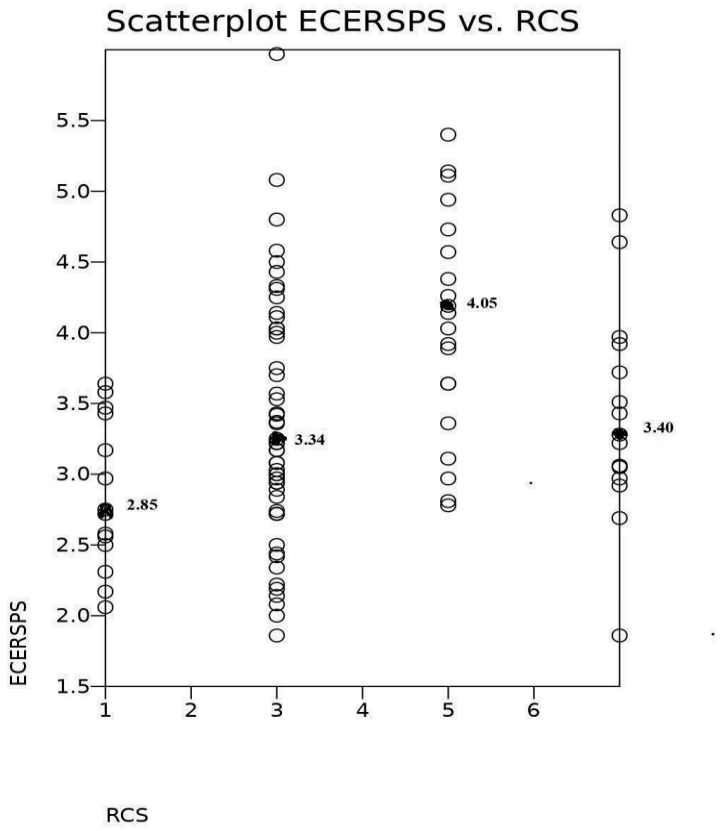
The first figure shows the actual individual violation data of the programs compared to their corresponding ECERS scores (SCATTERPLOT: ECERSPS vs LSTOTAL). There is not a significant relationship between the two as depicted in the graphic.



The following figure (SCATTERPLOT: ECERSPS vs RCS) 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 from 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 in the USA and Canada

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:

Table 2: 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 subtracted 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.

Table 3: 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.

Table 4: 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.

Table 5: ANOVAs Comparing the RCS Models with Program Quality

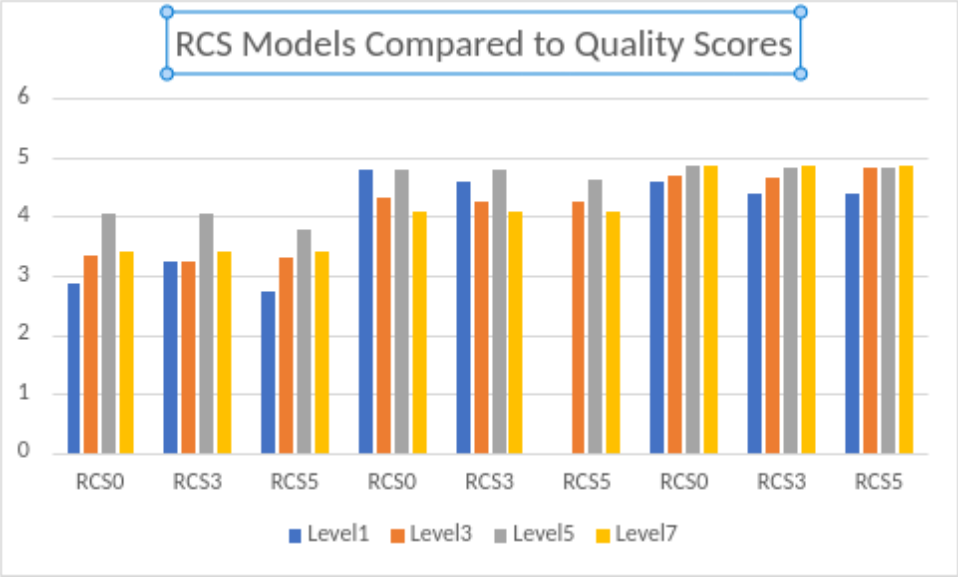
Jurisdictions	**Model**	**Level 1**	**Level 3**	**Level 5**	**Level 7**
Jurisdiction1	**RCS0**	**2.85**	**3.34**	**4.05**	**3.40**
	RCS3	3.24	3.23	4.05	3.40
	RCS5	2.73	3.32	3.77	3.40
Jurisdiction2	RCS0	4.81	4.31	4.80	4.10
	RCS3	**4.59**	**4.25**	**4.80**	**4.10**
	RCS5	---	4.26	4.64	4.10
Jurisdiction3	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 Regulatory Compliance 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).

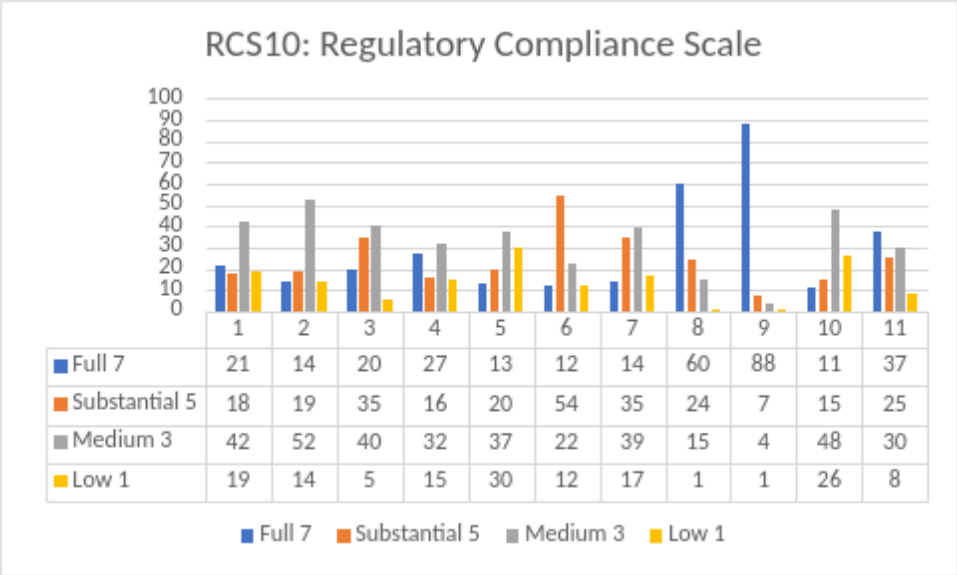
The RCS was based upon the following rubric: Full Regulatory Compliance (100%) or no violations = 7; Substantial Regulatory Compliance (99-98) or 1-2 violations = 5; Medium Regulatory Compliance (97-90) or 3-10 violations =3; and Low Regulatory Compliance (89 or less) or 11 or more violations = 1.

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).

Table 6			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.



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 for Quality Assessments

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 (Regulatory Compliance x Program Quality Grid Model) 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 mitigate 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.

Table 7: 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.

Table 8: 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 goes 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, the methodologies of key indicators, risk assessment and the resultant differential monitoring techniques have been used to determine 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 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 maintained 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, it has been found that certain jurisdictions have very unusually skewed data distributions which means that these ranges need to be tightened 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.

Regulatory Compliance Scale, Key Indicators, Risk Assessment, Differential Monitoring, and Program Quality

This section will expand the Regulatory Compliance Scale (RCS) conceptually to demonstrate how it relates to the key indicator and risk assessment methodologies as well as the infusion of program quality into rule development (Fiene, 2025a). The RCS has been pilot tested and demonstrated to be a viable alternative for measuring regulatory compliance with rules/regulations in the human services (Fiene, 2025b). The RCS moves regulatory compliance from a nominal based measurement strategy to an ordinal based measurement strategy. This change helps to enhance its statistical modeling capabilities which will mirror how more program quality systems operate: accreditation systems, such as the National Association for the Education of Young Children (NAEYC, 2025) and other program quality scales: the Environmental Rating Scales (Harms, Clifford, & Cryer , 2023). It also aligns more closely with the theory of regulatory compliance (Fiene, 2019, 2022, 2025a) in which a quality infusion component has been added to rule development and implementation.

To depict this relationship of regulatory compliance, key indicators, risk assessment, and program quality, the below table (Table 1: Regulatory Compliance Scale Plus) does a side-by-side comparison of these components. The first column shows how key indicators would play out at both the licensing and quality levels. The second column presents the regulatory compliance scale and shows how the infusion of quality builds upon full regulatory compliance. The third column shows how risk assessment interfaces with the regulatory compliance scale in which low risk rules (weights = 1-3) would be generally at a substantial compliance level with medium risk rules (weights = 4-6) being at a partial compliance level and lastly high-risk rules (weights = 7-9) being at a low compliance level. The fourth column suggests when differential monitoring (DM), in which targeted or abbreviated inspections are utilized focusing on key risk indicator rules, can be used in place of a comprehensive review (CR) when all rules are assessed. In studies (Fiene, 2025c), it has been demonstrated that key indicator rules predict either full or substantial compliance with all rules and are generally of a low overall risk; while risk assessment rules, especially those determined to be high risk rules, are usually always in compliance.

Recently, the Regulatory Compliance Scale and the risk assessment methodology have been combined with the Uncertainty-Certainty Matrix (UCM) in making licensing decisions (Fiene, 2025c). This combinatory effort has resulted in a more robust measurement strategy that helps to support moving from a nominal to ordinal measurement strategy. The UCM is used in determining the accuracy of each rule's regulatory compliance which enhances the reliability of the Regulatory Compliance Scale (RCSplus).

The purpose of this section was to demonstrate the interface amongst the various methodologies (key indicators and risk assessment) utilized within a differential monitoring approach in making licensing decisions (Fiene, 2025a), as well as showing how the key indicator methodology can be used for both licensing as well as quality indicator development (Fiene, 2025a).

Table 9: Regulatory Compliance Scale Plus (RCSplus)

Key Indicators	Regulatory Compliance Scale	Risk Assessment	Differential Monitoring
Quality Indicators	7+ = Exceeds Compliance		Yes
Licensing Indicators	7 = Full Compliance		Yes
Licensing Indicators	5 = Substantial Compliance	Low Risk (1-3)	Yes
	3 = Partial Compliance	Medium Risk (4-6)	No
	1 = Low Compliance	High Risk (7-9)	No

Conclusion

The Theory of Regulatory Compliance (Fiene, 2019, 2025a) 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, 2025c) 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.

Table 10: 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. In concluding, 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 final 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 (2022). 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 (2025a). Finding the Right Rules. ***American Scientist, Volume 113, 1***. pps 16-19.

Fiene (2025b). Potential Solution to the Child Care Trilemma Revisited – Finding the "Right Rules" – The Holy Grail of Early Care and Education***, Exchange***, Summer, 2025.

Fiene (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***.

Harms, Clifford, & Cryer (2023). ***Early Childhood Environmental Rating Scale (ECERS-3***),. https://ers.fpg.unc.edu/scales-early-childhood-environment-rating-scale-third-edition.html. Chapel Hill, North Carolina.

NARA (2023). ***Saskatchewan Differential Monitoring/Quality Indicators Scale Validation Study***, National Association for Regulatory Administration, Fredericksburg, Virginia.

NAEYC (2025). ***National Association for the Education of Young Children Accreditation System***. https://www.naeyc.org/accreditation. Washington, DC.

Richard Fiene PhD, Research Psychologist, Research Institute for Key Indicators (RIKILLC), Edna Bennett Pierce Prevention Research Center; Professor of Psychology (ret), Penn State University. Email: Rfiene@RIKInstitute.com Website: http://RIKInstitute.com

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

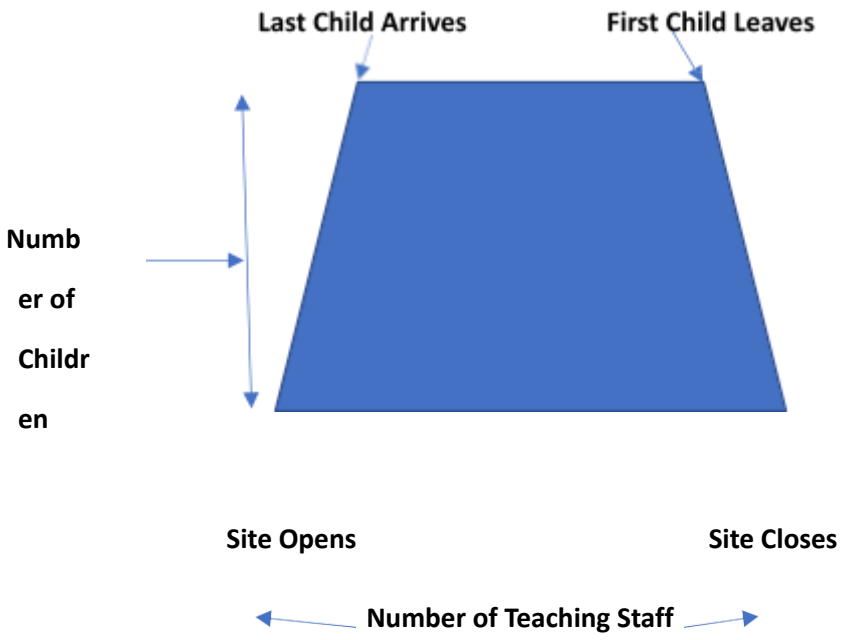
Group Size, Staff Child Ratio, Number of Children and Staff

<---------- 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.

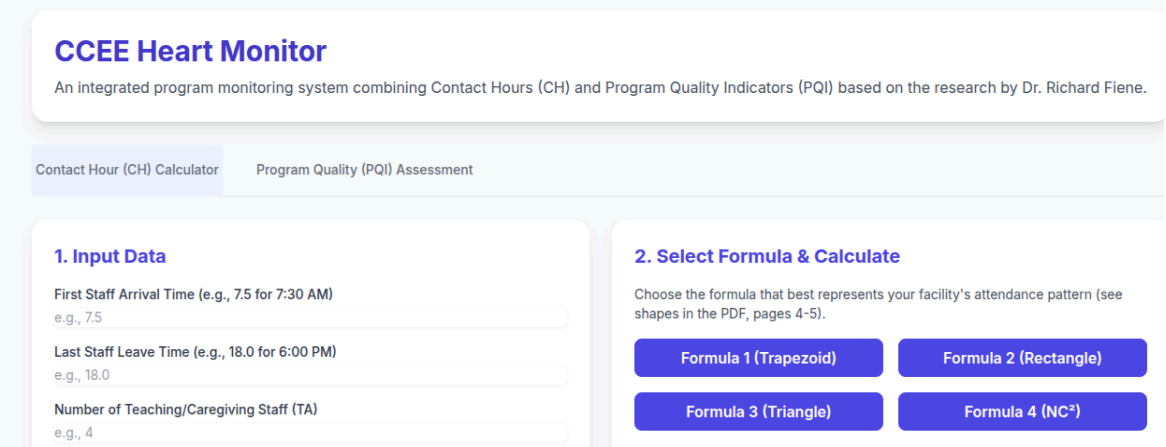
This scoring protocol (Table 2) which is generated from the Program Quality Indicators AI algorithms rank orders all programs on how well the classrooms measure up during the AI observations. The standardized scores are to the left while the actual classroom scores are to the right in the below table. 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/