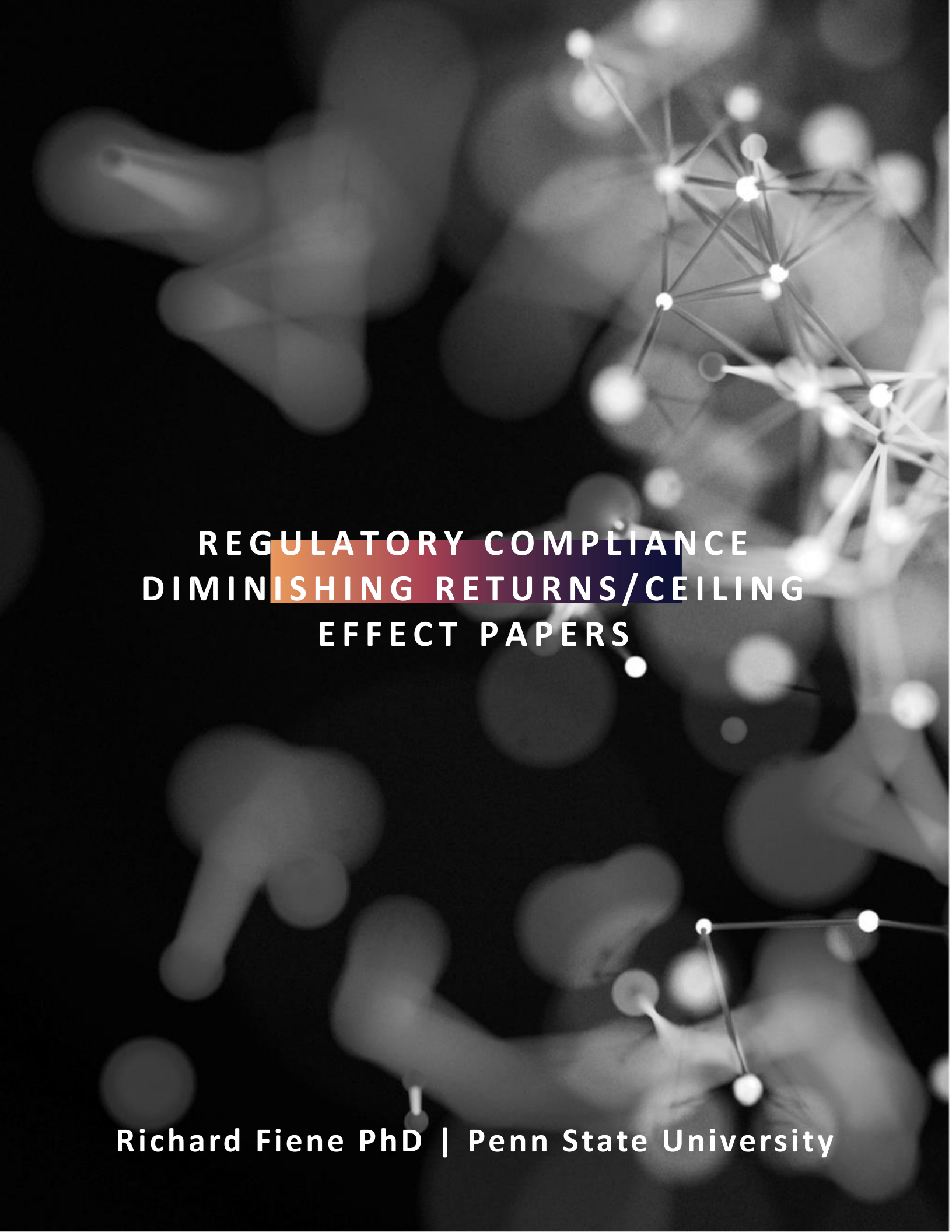


REGULATORY COMPLIANCE DIMINISHING RETURNS/CEILING EFFECT PAPERS

Richard Fiene PhD | Penn State University

The Regulatory Compliance Diminishing Returns/Ceiling Effect

Three documents describe the regulatory compliance diminishing returns/ceiling effect which has had an impact in the human services and potentially can have a large impact in other industries. These documents highlight and then provide the details to several studies conducted over the past decade that clearly depict this regulatory compliance diminishing returns/ceiling effect when comparing regulatory compliance and program quality in early care and education programs.

The purpose of this paper is to have regulatory scientists from other industries begin to think if the same regulatory compliance diminishing returns/ceiling effect occurs in their respective industry. It is clearly evident in the human services early care and education field as demonstrated by the studies included in this paper. This paper also provides possible mitigation strategies (Regulatory Compliance Scale and Quality Indicators) for dealing with this ceiling effect and they are interspersed throughout the three documents. It would be interesting to see if other industries have the same effect as in the human services and if some of the mitigation strategies proposed here would work in those other industries.

The three documents build upon each other with the first two providing an overview and more of a summary while the third provides the specific research reports documenting the studies done in the respective settings ranging from national in scope to state or provincial in scope.

The hope is that this paper generates a good deal of thinking and discussion around this discovery of a diminishing returns/ceiling effect with regulatory compliance as it relates to program quality.

Fiene, June, 2023

Introducing the Ceiling Effect/Diminishing Returns, Regulatory Compliance Scale, and the Quality Indicators Scale to Regulatory Science

Richard Fiene PhD

Research Institute for Key Indicators/Prevention Research Center/Penn State University

May 2023

The purpose of this short paper/public policy commentary is to introduce three relatively new, recently validated concepts to regulatory science. The first of the concepts (ceiling effect) is one that I have written about a good deal in previous policy commentaries when addressing the theory of regulatory compliance (Fiene, 2019). The other two (regulatory compliance and quality indicator scales (Fiene, 2022, 2023b; NARA, 2023)) have been validated more recently so they are relatively new, but I think will have a similar impact on the regulatory science field based upon the research interest generated worldwide.

The "Ceiling Effect" is a more user-friendly term for the theory of regulatory compliance diminishing returns. I have found in recent webinars and presentations that the notion of a ceiling effect resonates with other regulatory science researchers more so than the theory of regulatory compliance diminishing returns. Scientists can wrap their heads around the ceiling effect much easier than the theory, so I am going to use this new term rather than the older. However, they do mean the same thing, same result, just different terminology. It is similar to what happened with "inferential inspections" (earlier term) and "differential monitoring" (present terminology) (Fiene, 2023a). Same concept, just different terms.

The "ceiling effect" is the same relationship between regulatory compliance and program quality. As regulatory compliance increases from substantial compliance to full 100% compliance, program quality shows either no improvement or diminished improvement over the same course. This is the essence of the theory of regulatory compliance diminishing returns (Fiene, 2019, 2023a, 2023b; NARA, 2023). No change here.

The second concept I want to introduce is 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.

The last concept to be introduced deals with quality indicators which have been proposed as part of a differential monitoring paradigm but not utilized and validated in specific jurisdictions. Well, that has changed now with a major study completed in the Province of Saskatchewan which has clearly demonstrated in a valid and reliable fashion how quality indicators can be used effectively and efficiently when compared to other program quality scales and regulatory compliance data (NARA, 2023).

All these above results (Fiene, 2023b; NARA, 2023) were part of this Province of Saskatchewan five-year project, and they are all in the early care and education domain, but I think that the results are pertinent to any industry governed by regulatory science principles. One needs to change the content obviously, but the metrics and methodology would hold up because of their base in solid scientific principles of instrument and research design.

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 Scale, ***RIKINotes Blog***, January 2022.

Fiene (2023a). ***Licensing Measurement & Monitoring Systems***, Research Institute for Key Indicators, Elizabethtown, Pennsylvania.

Fiene (2023b). Ceiling Effect/Diminishing Returns, Regulatory Compliance Scale, and Quality Indicators Scale, ***Mendeley Data,*** doi: 10.17632/gc423hprcs.1

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

The Public Policy Implications of the Regulatory Compliance Theory of Diminishing Returns, Regulatory Compliance Scaling, and the Program Quality Scoring Matrix along with Integrative Monitoring

Richard Fiene PhD

Research Institute for Key Indicators/Penn State

March 2023

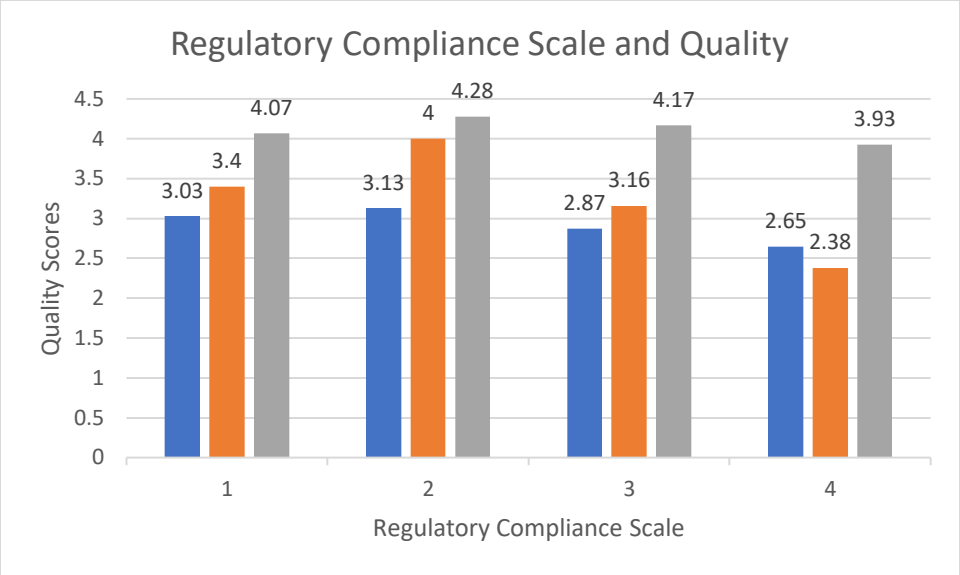
This technical research note/abstract provides a data matrix (below table) depicting the relationship between regulatory compliance and program quality. The data clearly demonstrate the regulatory compliance theory of diminishing returns which depicts the ceiling or plateau effect in this relationship between regulatory compliance data and program quality data. It also shows the difficulty one will have in distinguishing program quality differences at the full and high regulatory compliance levels but the ease in distinguishing program quality between low regulatory compliance and high regulatory compliance levels.

This abstract unifies several separately developed regulatory compliance metrics and concepts by combining them into a single technical research note. The Regulatory Compliance Theory of Diminishing Returns (2019), The Regulatory Compliance Scale (2022), Integrative Monitoring (2023), and the Ten Principles of Regulatory Compliance Measurement (2023) have all been presented separately (all these papers are available for the interested reader on ***SSRN (***https://www.ssrn.com/index.cfm/en/***) or the Journal of Regulatory Science (***https://regsci-ojs-tamu.tdl.org/regsci/***)***). This abstract shows how they are all related and their importance in moving forward with regulatory compliance measurement in the future. The four jurisdiction's (US National, Southern State, Western State, Canada) final reports are available at https://www.naralicensing.org/key-indicators for the interested reader.

Relationship of Regulatory Compliance Scale and Program Quality in Four Jurisdictions Matrix

Reg Comp Scale	US National	Southern State	Western State	Canada
Full	3.03 (75)	3.40 (15)	4.07 (82)	37.4 (44)
High	3.13 (135)	4.00 (20)	4.28 (69)	38.5 (33)
Mid	2.87 (143)	3.16 (32)	4.17 (163)	29.1 (36)
Low	2.65 (28)	2.38 (2)	3.93 (71)	----------
Significance	***p < .001***	***p < .05***	***p < .001***	***p < .01***

Legend:
US National = CLASS-IS scores
Southern State and Western State = ECERS-R scores
Canada = Canadian Program Quality Tool scores
One-way ANOVA was performed on the data in each jurisdiction.
Regulatory Compliance Scale (Reg Comp Scale (RCS)):
Full = 0 violations (100% regulatory compliance with all rules/regulations)
High = 1-2 violations
Mid = 3-9 violations
Low = 10+ violations
The number in parentheses is the number of programs assessed in each jurisdiction.



Legend:
1 = Full; 2 = High; 3 = Mid; 4 = Low.
Blue = US National; Orange = Southern State; Gray = Western State. Canada was left off because of different scaling.

The above data matrix display is important for the early care and education (ECE) field because it demonstrates the relationship between licensing via regulatory compliance data measurement and program quality scores via CLASS, ERS, and the Canadian Quality Tool. The CLASS and ERS are well grounded ECE program quality tools while the Canadian Quality Tool is a new addition to the field.

The data displayed show that a ceiling or plateau effect (quality scores did not change significantly as was generally the case with lower levels of regulatory compliance) occurred in all four jurisdictions when the regulatory compliance levels or the absence of rule/regulatory violations were compared to program quality scores as one moves from high regulatory compliance to full regulatory compliance (0 violations or 100% regulatory compliance with all rules). From a public policy point of view, it would lead us to believe that licensing is not the best avenue to program quality and that another intervention, such as Quality Rating and Improvement Systems (QRIS), would be necessary to enhance quality programming. What regulatory compliance and licensing does do is prevent harm and keep children in healthy and safe environments (please go to https://rikinstitute.com for examples to support this claim). So, from a public policy point of view, licensing is accomplishing its goals. But don't expect licensing to address quality programming. For that to occur, either we need to continue our present system of licensing and Quality Initiatives, such as QRIS, as an add on; or infuse quality into the rules and regulations which has been suggested via a new form program monitoring called: integrative monitoring.

There are some other takeaways from the above data matrix that are significant contributions to the regulatory compliance measurement research literature, such as, how skewed the data are. Focus more on the number of programs rather than their quality scores for each of the Regulatory Compliance Scale levels. You will notice that most programs in each of the jurisdictions are either in full or high regulatory compliance and that there are few programs at the low end of the regulatory compliance scale. There is an unusually very high percentage of programs at full compliance. This also contributes to a lack of

variance in the upper end of the regulatory compliance scale which can be problematic as indicated in the previous paragraph in distinguishing between the quality levels of programs.

The importance of these four studies and the summary matrix above is to provide a context in how licensing and regulatory compliance data should be used in making public policy decisions, for example: is it more effective and efficient to require high or substantial regulatory compliance than full regulatory compliance with all rules and regulations to be granted a full license to operate? It appears prudent to continue with the US emphasis on QRIS as an add on quality initiative, especially in states where rules/regulations are at a minimal level. In Canada their emphasis has been more in line with an integrative monitoring approach in which quality elements are built in or infused within the rules and regulations themselves. This approach appears to work in a similar fashion and is an effective public policy initiative. Either approach appears to be an effective modality to increasing program quality; but are both equally efficient.

Richard Fiene PhD, Research Psychologist and Regulatory Scientist, Research Institute for Key Indicators and Penn State University; Senior Research Consultant, National Association for Regulatory Administration. Contact Dr Fiene at rfiene@rikinstitute.com or rfiene@naralicensing.org if interested in learning more about the regulatory compliance studies.

Five Studies Providing the Empirical Evidence for the Regulatory Compliance Diminishing Returns Effect: Additional Support for the Theory of Regulatory Compliance

May 2023

The attached five studies, three from the United States and two from Canada, provide the empirical evidence supporting the regulatory compliance diminishing returns effect or ceiling effect. These studies clearly demonstrate that moving from substantial regulatory compliance to full (100%) regulatory compliance correlates with a diminishing returns effect or ceiling effect in the specific program quality measures, such as the CLASS, ECERS, FDCRS, ITERS, a provincial program quality tool and the Fiene Program Quality Indicators Scale. In other words, the fully compliant programs' quality scores were either not significantly different from the substantially compliant programs or there was an actual drop off in the quality scores.

These combined results have significant implications for public policy in the early care and education arena in which a more targeted type of public policy regarding the right rules to be complied with, rather than all rules to be complied with, may be in the best interests of program quality.

Fiene, 2023

Georgia Child Care Licensing Study: Validating the Core Rule Differential Monitoring System

Executive Summary

Richard Fiene, Ph.D.

The purpose of this study was to validate Georgia's process for determining if a state-regulated child care facility is compliant with basic state health and safety requirements. The process was developed by staff at Bright from the Start: Georgia Department of Early Care and Learning (DECAL). Currently Georgia utilizes a "Core Rule" risk assessment approach in which the health and safety rules deemed most crucial to ensure children's health and safety are used to compute a program's compliance status.

This validation study utilized a unique analytical model that compared licensing data with previous key indicator (*for readers not familiar with this term, please see the definitions on page 4 of the report*) research and ascertained if the Core Rules accurately indicated a program's overall compliance with the total population of licensing rules.

Additional statistical analyses examined if the mathematical formula used to compute compliance was an appropriate configuration of the data that discerned between those programs that adequately met basic health and safety rules (compliant) and those that did not (non-compliant). Also licensing data were compared to a representative sample of quality data collected as part of a different study to examine the correlation between compliance and quality. A Differential Monitoring Logic Model/Algorithm (*DMLMA©*) (Fiene, 2012) and a previous validation model (Zellman & Fiene, 2012) were used in the research.

One hundred and four child care centers (104 CCC) and 147 family child care (FCC) homes were assessed. Licensing data over a four-year period (2008-2012) and matching program quality data from a two-year period (2007-2008) were used in this study.

The study focused on three research questions:

1. Do the Core Rules CCCs and FCC homes serve as overall Key Indicators of compliance?
2. Does the Annual Compliance Determination Worksheet (ACDW) appropriately designate programs as compliant or non-compliant related to health and safety?
3. Are the Core Rules related to program quality?

The analysis demonstrated that the Core Rules did serve as key indicators, and these key indicators were identified for both center based and home based child care. The second analysis concluded that the ACDW computation did distinguish between compliant and non-compliant programs. Finally, the expected correlation between compliance and quality was found but only for state-funded Pre-K classrooms, not for family child care nor for preschool classrooms that were not part of the state-funded Pre-K.

Georgia Child Care Licensing Study: Validating the Core Rule Differential Monitoring System

Richard Fiene, Ph.D.

February 1, 2014

This study was made possible by a grant from Bright from the Start: Georgia Department of Early Care and Learning. All opinions expressed in the report reflect the opinions of the author, not necessarily those of the Department of Early Care and Learning.

ABSTRACT

The purpose of this study was to validate Georgia's process for determining if a state-regulated child care facility is compliant with basic state health and safety requirements. The process was developed by staff at Bright from the Start: Georgia Department of Early Care and Learning (DECAL). Currently Georgia utilizes a "Core Rule" risk assessment approach in which the health and safety rules deemed most crucial to ensure children's health and safety are used to compute a program's compliance status. This validation study utilized a unique analytical model that compared licensing data with previous key indicator (*for readers not familiar with this term, please see the definitions on page 4 of the report*) research and ascertained if the Core Rules accurately indicated a program's overall compliance with the total population of licensing rules. Additional statistical analyses examined if the mathematical formula used to compute compliance was an appropriate configuration of the data that discerned between those programs that adequately met basic health and safety rules (compliant) and those that did not (non-compliant). Also licensing data were compared to a representative sample of quality data collected as part of a different study to examine the correlation between compliance and quality. A Differential Monitoring Logic Model/Algorithm (*DMLMA*©) (Fiene, 2012) and a previous validation model (Zellman & Fiene, 2012) were used in the research. Child care centers (CCC) and family child care (FCC) homes were assessed. The analysis demonstrated that the Core Rules did serve as key indicators, though this list should be reexamined. The second analysis concluded that the computation could be simplified. Finally, the expected correlation between compliance and quality was found but only in state-funded Pre-K classrooms; it was not found in preschool classrooms and could not be validated. Family child care could not be validated either. As a result of the study, recommendations were made to strengthen Georgia's system.

Acknowledgements:

Special thanks are extended to DECAL staff who had the vision to conduct this validation study: Bobby Cagle, Commissioner; Kay Hellwig, Assistant Commissioner for Child Care Services; Kristie Lewis, Director of Child Care Services; and Dr. Bentley Ponder, Director of Research & Evaluation. Also, researchers at the University of North Carolina, Chapel Hill, Frank Porter Graham Child Development Institute , Dr. Donna Bryant and Dr. Kelly Maxwell who made this study so much more significant by sharing program quality data from earlier studies they completed in Georgia.

INTRODUCTION

Background of Georgia's Compliance Determination System

Similar to other states, Georgia has a licensing and monitoring system that oversees a diverse population of early care and learning programs across the state. The licensing and monitoring system of early care and learning programs is charged to Bright from the Start: Georgia Department of Early Care and Learning (DECAL), a state early education department that also oversees and administers Georgia's Pre-K Program, Child Care and Development Block Grant, the Child and Adult Care Food Program, and the Summer Food Service Program. In 2012, DECAL's licensing and monitoring system regulated approximately 6,300 early care and learning programs. The crux of this regulation is determining if the programs meet Georgia's health and safety rules. Programs that meet these rules are determined to be compliant.

In the mid 2000's, Georgia began experimenting with a process that determined whether or not a program was designated as compliant with the state's health and safety regulations by focusing on key Core Rules. These are health and safety rules deemed crucial to minimizing risk related to children's health and safety. Seventy-four rules out of the 456 that programs must follow were classified as Core Rules[1]. Core Rules are cited by severity (low, medium, high, extreme). It is important to note that this entails a risk assessment theoretical approach rather than a Key Indicator statistical approach. This means that the Core Rules were determined by content analysis rather than by a statistical procedure.

Though this system has undergone some slight revisions, this basic methodology is still in place:

1. All programs receive at least one full licensing study and one monitoring visit. At the licensing study all applicable rules are examined. At the monitoring visit, only Core Rules (or any rule that was not met at the licensing study) are examined.
2. If additional visits are conducted, the Core Rules are examined again at that time.
3. At the end of the fiscal year (June 30), each program receives a compliance determination. This determination is based on all visits (licensing study, monitoring visit, and other reviews). A standardized worksheet, Annual Compliance Determination Worksheet (ACDW), is used to make the computation that determines the designation.
4. The compliance status remains until the next determination one year later. Programs do not have an opportunity to contest the compliance determination, though programs have numerous opportunities to contest any citation.
5. At the conclusion of Fiscal Year 2012, approximately 91% of the programs were classified as compliant. A program's eligibility for certain services, acceptance into Quality Rated and Georgia's Pre-K Program, is impacted by the program's compliance determination.

Background of this Study

Since the compliance determination system has been used for several years, key policymakers at DECAL requested an external review to validate if the system was operating as intended. Are the Core Rules a sufficient subsample to measure a program's overall regulation with the state's health and safety regulations? Furthermore, does the compliance determination formula appropriately differentiate compliant programs from non-compliant programs? In other words, is the computation a viable way to make this designation? And finally, does compliance determination serve as a sufficient indicator for other aspects of quality not addressed in Georgia's health and safety rules?

The purpose of this study was to validate the aforementioned compliance determination process. This validation process utilized a unique analytical model that compared licensing data with previous key indicator research and ascertained if the Core Rules are an indication of a program's overall compliance with the total population of licensing rules. Second, additional statistical analyses examined if the mathematical formula used to compute compliance was an appropriate configuration of the data that differentiated between those programs that adequately met basic health and safety rules (compliant) and those that did not (non-compliant). Finally, licensing data were

[1] The number of Core Rules was expanded in 2012 to include increased enforcement and sanctions regarding transportation. The new Core Rules were not part of this analysis.

compared to a representative sample of quality data collected as part of a different study to examine the correlation between compliance and quality (see a further explanation of the sample in the Limitations Section of this report).

Specifically, the study addressed the following research questions:

1. **Do the Core Rules for child care centers (CCC) and family child care (FCC) homes serve as overall Key Indicators of compliance?**
2. **Does the Annual Compliance Determination Worksheet (ACDW) appropriately designate programs as compliant or non-compliant related to health and safety?**
3. **Are the Core Rules related to program quality?**

The following definitions are used in the study:

Core Rules = the rules determined to be of greatest importance and place children at greatest risk if not complied with. This approach is defined in the licensing literature as a risk assessment approach. Core Rules cover 12 regulatory areas and 74 specific rules. The Core Rules were the focal point of this validation study and are addressed in the first approach to validation – Standards and the first research question.

ACDW = Annual Compliance Determination Worksheet, the compliance decision-making system based on the Core Rules that can be used to determine the number of visits made to programs. The ACDW was the secondary focal point of this validation study and is addressed in the second approach to validation – Measures and the second research question.

Key Indicators = a differential monitoring approach that uses only those rules that statistically predict overall compliance with all the rules. In other words, if a program is 100% in compliance with the Key Indicators, the program will also be in substantial to full compliance with all rules. The reverse is also true in that if a program is not 100% in compliance with the Key Indicators, the program will also have other areas of non-compliance with all the rules. In this study, eight Key Indicators rules were identified for CCC and nine Key Indicators rules for FCC (See pages 15-16 for the specific indicators and additional detail about the methodology). These are in addition to the Core Rules.

Rule Violations or Citations = occurs when a program does not meet a specific rule and is cited as being out of compliance with that rule. These individual rule violations/citations are summed to come up with total violation/citation scores on the Core Rules and on the Licensing Studies.

Differential Monitoring = a relatively new approach to determining the number of licensing visits made to programs and to what rules are reviewed during these visits. Two measurement tools drive differential monitoring: one is a Weighted Risk Assessment, and the other is a Key Indicator checklist. Weighted Risk Assessments determine how often a program will be visited while Key Indicator checklists determine what rules will be reviewed in the program. Differential monitoring is a powerful approach when Risk Assessment is combined with Key Indicators because a program is reviewed by the most critical rules and the most predictive rules. See Figure 1 which presents a Logic Model & Algorithm for Differential Monitoring *(DMLMA©) (*Fiene, 2012).

Licensing Study = a comprehensive review of a program where all child care rules are reviewed.

Monitoring Visit = an abbreviated form of a visit and review in which only a select group (Core Rules) of child care rules are reviewed.

Program Quality = for the purposes of this study, quality was measured in child care centers by the *Early Childhood Environment Rating Scale-Revised (ECERS-R), Infant Toddler Environment Rating Scale-Revised (ITERS-R*) and in family child care homes by the *Family Child Care Environment Rating Scale-Revised (FCCERS-R*). The program quality measures were used as part of the third approach to validation – Outputs and the third research question.

Scoring for Licensing Variables/Data Collection Protocols:

Licensing Study = the total number of rule violations for a specific facility.

Core Rules = the total number of core rule violations.

ACDW/Compliance Designation = the annual compliance determination taken from the Annual Compliance Determination Worksheet. Compliant [C] was coded as "1" in the data base; Non-Compliant [NC] was coded as "0" in the data base.

Key Indicators = these were generated by a statistical methodology based upon the ability of the specific rule to predict full compliance with all the rules. Data from the Licensing Studies were used to make this determination of key indicator rule status.

METHODOLOGY AND ANALYTICAL FRAMEWORK

Licensing data over a four-year period (2008-2012) and matching program quality data from a two-year period (2007-2008) were used in this study. Specifically, data from 104 child care centers and 147 family child care homes were analyzed. Data from licensing studies (all rules) and monitoring visits (selected rules) were utilized. Program quality data were provided by researchers from the FPG Child Development Institute at the University of North Carolina at Chapel Hill (FPG), and the FPG research team matched these data points with the licensing data provided by DECAL (See the following website for the specific reports - ***http://decal.ga.gov/BftS/ResearchStudyOfQuality.aspx)***. All the data were analyzed by the Research Institute for Key Indicators.

Two models were used to frame the analysis: a Validation Framework that uses four approaches (Zellman & Fiene, 2012) to validating quality rating and improvement systems (QRIS) being applied to licensing systems; and a *Differential Monitoring Logic Model and Algorithm (DMLMA©)(*Fiene, 2012) were employed to answer the three research questions for this Validation Study. The validation approaches are described below; the *DMLMA*© is described at the beginning of the Findings Section of this report.

The first validation approach deals with examining the validity of key underlying concepts by assessing if basic components and standards are the right ones by examining levels of empirical and expert support. For this study, this approach used Key Indicators to validate the Core Rules since Risk Assessment and Key Indicators are differential monitoring approaches. This answers the first research question.

The second validation approach deals with examining the measurement strategy and the psychometric properties of the measures used by assessing whether the verification process for each rule is yielding accurate results. Properties of the key rules can be measured through inter-rater reliability on observational measures, scoring of documentation, and inter-item correlations to determine if measures are psychometrically sound. Cut scores can be examined to determine the most appropriate ways to combine measures into summary ratings. For this study, the second validation approach validates the use of the ACDW and Core Rules by comparing compliance decisions with the Licensing Studies. This answers the second research question.

The third validation approach deals with assessing the outputs of the licensing process by examining the variation and patterns of program level ratings within and across program types to ensure that the ratings are functioning as intended. The approach examines the relationship of program level ratings to other more broadly based program quality measures and examines alternate cut points and rules to determine how well the ratings distinguish different levels of quality. For this study, this approach used data from Core Rules and Licensing Studies and data from earlier program quality studies (Maxwell, et al., 2009a,b; 2010) for validation. This answers the third research question.

Out of the four validation approaches (See Table 8), only three were utilized in this study. ***The fourth validation approach*** deals with how ratings are associated with children's outcomes. This approach examines the relationship

between program level ratings and selected child outcomes to determine whether higher program ratings are associated with better child outcomes. This approach did not have data that could be used in this study.

FINDINGS

The *DMLMA*© (See Figure 1) provides the conceptual model for assessing the overall effectiveness of Georgia's approach using Core Rules. In the model, the two main tools are Risk Assessment and Key Indicator measurements, which are created from a statistical analysis of the comprehensive licensing tool. The comprehensive licensing tool measures compliance with all rules. For the purposes of this study the Licensing Study represents the comprehensive licensing tool while the Core Rules represent a Risk Assessment tool. For the Program Quality tools, the ECERS-R, ITERS-R and FCCERS-R were utilized from an earlier program quality study by FPG Child Development Institute at the University of North Carolina at Chapel Hill (Maxwell, et al., 2009a,b; 2010). Georgia currently does not use a Key Indicator tool (see Table 1). With the DMLMA© analytical methodology, specific correlational thresholds are expected (please refer to Figure 1 on page 14).

TABLE 1

***DMLMA*© Terminology**	**Georgia Examples and Data Sources**
Comprehensive Tool	Licensing Study
Program Quality Tool	ECERS-R and ITERS-R for CCC; FCCERS-R for FCC
Risk Assessment Tool	Core Rules
Key Indicators Tool	Not Present (Generated as part of this Study-see Tables 9/10)
Differential Monitoring Tool	ACDW Compliance Determination

Before presenting the findings for the validation approaches, some basic descriptive statistics are provided regarding the major variables in this study: Licensing Study, ACDW, Core Rules, and Key Indicators (see Table 2). The data are provided for both child care centers and family child care homes. It is clear from these basic descriptive statistics that the data distributions are very skewed in a positive fashion which means that there is very high compliance with all the major licensing variables for this study. In other words, the majority of programs are in substantial compliance with all the licensing rules and receive a compliant determination.

TABLE 2

Licensing Variable	**Mean**	**Range**	**SD**	**Skewness**	**Kurtosis**
Licensing Study (CCC)	5.51	25	5.26	1.47	2.11
ACDW (CCC)	0.75	1	0.44	-1.17	-0.64
Core Rules (CCC)	4.47	22	4.72	1.81	3.60
Key Indicators (CCC)	1.68	6	1.61	0.90	0.073
Licensing Study (FCC)	5.85	33	5.71	1.56	3.37
ACDW (FCC)	0.87	1	0.34	-2.23	3.03
Core Rules (FCC)	1.61	11	1.75	1.99	6.61
Key Indicators (FCC)	2.37	8	2.13	0.63	-0.57

Licensing Study Mean = the average number of total rule violations.
ACDW Mean = the average score for a determination of compliance (1) or non-compliance (0).
Core Rules Mean = the average number of core rule violations.
Key Indicators Mean = the average number of key indicator violations.

The findings are presented by the three validation approaches of Standards, Measures, and Outputs as well as the three research questions related to Key Indicators, Core Rules, and Program Quality.

1) Validation of Standards (First Approach to Validation) for answering the first research question: *Do the Core Rules for child care centers (CCC) and family child care (FCC) homes serve as overall key indicators of compliance?*

In this first approach to validation which focuses on Standards, Key Indicators were generated from the Licensing Studies because Core Rules (a Risk Assessment tool) and Key Indicators are both Differential Monitoring approaches (see Figure 1). The Core Rules were compared to the Key Indicators generated by the licensing data base and there was a .49 correlation for CCC (n = 104) and .57 correlation for FCC (n = 147) which indicates a

relationship between the Core Rules and Key Indicators at a $p < .0001$ significance level (Table 3). Also, the Key Indicators were correlated with the Licensing Study data and significant results were determined with r values of .78 ($p < .0001$) for CCC ($n = 104$) and .87 ($p < .0001$) for FCC ($n = 147$). These results clearly met the expected *DMLMA*© thresholds between the key indicator rules with core rules (.50+) and licensing studies (.70+).

TABLE 3

Key Indicators with Core Rules and Licensing Study	**r =**	**p <**	**n =**
Key Indicators and Core Rules (CCC)	.49	.0001	104
Key Indicators and Licensing Study (CCC)	.78	.0001	104
Key Indicators and Core Rules (FCC)	.57	.0001	147
Key Indicators and Licensing Study (FCC)	.87	.0001	147

Table 3 begins to demonstrate how the Georgia Child Care Licensing system is utilizing the ***DMLMA*©** terminology from Table 1. With the generation of Key Indicators from this study, all the key elements within a differential monitoring system are present. This crosswalk to the ***DMLMA*©** will continue in Tables 4 & 5.

2) Validation of Measures (Second Approach to Validation) for answering the second research question*: Is the Annual Compliance Determination Worksheet (ACDW) a valid measure in determining the overall health and safety compliance of Georgia's early care and learning programs?*

The Core Rules and the ACDW were compared to the Licensing Study data and compliance designation to determine the validation of the ACDW scoring protocol. There was a high correlation between the number of violations on the Core Rules and the total licensing violations on the Licensing Studies ($r = .69$; $p < .0001$)(Table 4). This result helps to validate that the ACDW is actually discriminating between high compliant and low compliant providers for CCC. For FCC, there was also a high correlation between the number of violations on the Core Rules and the total licensing violations on the Licensing Studies ($r = .74$; $p < .0001$). These results meet the *DMLMA*© thresholds of .50+ for Licensing Studies and Core Rules.

When Core Rules were correlated with the ACDW compliance decisions, there was a significantly high correlation for CCC ($r = .76$; $p < .0001$) and for FCC ($r = .70$; $p < .0001$). The key element of the ACDW scoring protocol is that the Core Rules distinguish between high and low compliant providers. The CCC/Core Rules and ACDW have been validated, as well as the FCC/Core Rules and ACDW because both the correlations were above the expected *DMLMA*© threshold (.50+).

TABLE 4

Core Rules with Licensing Studies and ACDW	**r =**	**p <**	**n =**
Core Rules and Licensing Studies (CCC)	.69	.0001	104
Core Rules and ACDW (CCC)	.76	.0001	104
Core Rules and Licensing Studies (FCC)	.74	.0001	147
Core Rules and ACDW (FCC)	.70	.0001	147

3) Validation of Outputs (Third Approach to Validation) for answering the third research question: *Are the Core Rules correlated with program quality?*

For this approach, programs were divided into those that had an ITERS-R score, an ECERS-R score for a preschool class, and an ECERS-R score for a Georgia's Pre-K class; and those that had only an ITERS-R score and an ECERS-R score for preschool. The sample was evenly divided. Since Georgia has placed substantial resources into its Pre-K program, it was thought that this analysis might suggest if there was anything different between programs with a Georgia's Pre-K class and those without.

When the Core Rules for CCC's were compared with program quality data (ECERS-R/PS + ITERS-R), a significant correlation was not found between CCC ($r = .27$) for programs with only preschool classrooms but was found for programs with Pre-K classrooms (ECERS-R/PK + ITERS-R) ($r = .60$). When Core Rules for FCC's were compared

to the FCC program quality data (FCCERS-R), the correlations were at a much lower level (r = .17) (See Table 5). However, these results are constrained by the limited range of the data; see the Limitation Section that follows this section.

Upon closer inspection of the correlations in Table 5 for CCC, it would appear that the CCC compliance system is more valid with the state-funded Pre-K programs (.48) than with the preschool programs (.21) because the correlations between the various Environment Rating Scales (ECERS-R + ITERS-R) are significant only when compared to the respective compliance with all rules on the Licensing Studies in the programs that have Pre-K programs. In making these comparisons, programs that had both ECERS-R and ITERS-R were combined and compared to the respective Licensing Study data (these data were reversed scored in which the number of violations were subtracted from a perfect score of 100). The differences are even more significant when you compare the Environment Rating Scales and the Core Rules where the Pre-K programs' correlation between the compliance with Core Rules and Environment Rating Scales is .60 and preschool programs is .27 while the FCC is .17.

Program quality data refer to data collected in earlier studies by researchers from FPG (Maxwell, et al., 2009a,b; 2010) in which FPG collected Environment Rating Scales (ECERS-R; ITERS-R; FCCERS-R) data on a representative sample of CCC and FCC (See (h***ttp://decal.ga.gov/BftS/ResearchStudyOfQuality.aspx)*.** In comparing the program compliance and program quality data, the analyses supported the validation of the CCC for Pre-K only programs (*DMLMA*© threshold = .30+) but it was weaker for the FCC programs and not significant for preschool programs and therefore could not be validated. See Table 13 on page 17 for a further explanation of the CCC data distribution.

TABLE 5

Program Compliance and Quality Comparisons	**r =**	**p <**	**n=**
ECERS-R/PK + ITERS-R and Licensing Studies	.48	.001	45
ECERS-R/PK + ITERS-R and Core Rules	.60	.0001	45
ECERS-R/PS + ITERS-R and Licensing Studies	.21	ns	45
ECERS-R/PS + ITERS-Rand Core Rules	.27	ns	45
FCCERS-R and Licensing Studies	.19	.04	146
FCCERS-R and Core Rules	.17	.03	146

LIMITATION

The sampling for this study was based on previous studies (Maxwell, 2009a,b; 2010) completed by FPG in which program quality data were collected and analyzed. This study employed a subset of sites that were a representative sample of Georgia's child care licensing system. Not all of these sites could be used for this study because some had closed or some did not have the necessary data to make comparisons. So the sample at this point is one of convenience; however, 104 of the 173 CCC and 146 of the 155 FCC were used in this study, a significant number of the original representative sample. Also, when the Environment Rating Scales (ECERS-R, ITERS-R, FCCERS-R) scores were compared with the CCC and FCC samples, there were no significant differences (average difference was .01-.03) between the two study samples (See Table 6).

TABLE 6

Environment Rating Scale Scores	**FPG**	**This Study**
ECERS-R Pre-K Total Scale Scores	4.16	4.15
ECERS-R Preschool Total Scale Scores	3.39	3.42
ITERS-R Total Scale Scores	2.74	2.72
FCCERS-R Total Scale Scores	2.50	2.49

CONCLUSION

The CCC differential monitoring through the Core Rules/ACDW has been validated on the three approaches (Standards, Measures, and Outputs (Pre-K Program only)) and three research questions (Key Indicators, Core Rules, Program Quality (Programs with Georgia Pre-K only)) (See Table 7). The FCC differential monitoring through the Core Rules/ACDW was validated on the first validation approach (Standards) and first research question (Key Indicators); validated on the second validation approach (Measures) and second research question (Core Rules); but not validated on the third validation approach (Outputs) and third research question (Program Quality).

TABLE 7

	Correlations		
Validation Approach/Research Question	**CCC Actual (Expected*)**		**FCC Actual (Expected)**
1 STANDARDS/Key Indicators	VALIDATED		VALIDATED
Key Indicators x Core Rules	.49 (.50+)		.57 (.50+)
Key Indicators x Licensing Studies	.78 (.70+)		.87 (.70+)
2 MEASURES/Core Rules/ACDW[2]	VALIDATED		VALIDATED
Core Rules x Licensing Studies	.69 (.50+)		.74 (.50+)
Core Rules x ACDW	.76 (.50+)		.70 (.50+)
3 OUTPUTS/Program Quality	VALIDATED		NOT VALIDATED
Licensing Studies x ERS**/PK	.48 (.30+)	FCCERS	.19 (.30+)
Core Rules x ERS/PK	.60 (.30+)	FCCERS	.17 (.30+)
Licensing Studies x ERS/PS	----------		.21 (.30+)
Core Rules x ERS/PS	----------		.27 (.30+)

DMLMA© Expected r Value Thresholds in Order to be Validated (Also see Figure 1 for additional details):
High correlations (.70+) = Licensing Studies x Key Indicators.
Moderate correlations (.50+) = Licensing Studies x Core Rules; Core Rules x ACDW; Core Rules x Key Indicators; Key Indicators x ACDW.
Lower correlations (.30+) = Program Quality Tools x Licensing Studies; Program Quality x Core Rules; Program Quality x Key Indicators.

Program Quality Tools = ECERS-R, ITERS-R, FCCERS-R.

****ERS = ECERS-R + ITERS-R**
PK = Pre-K program
PS= Preschool program

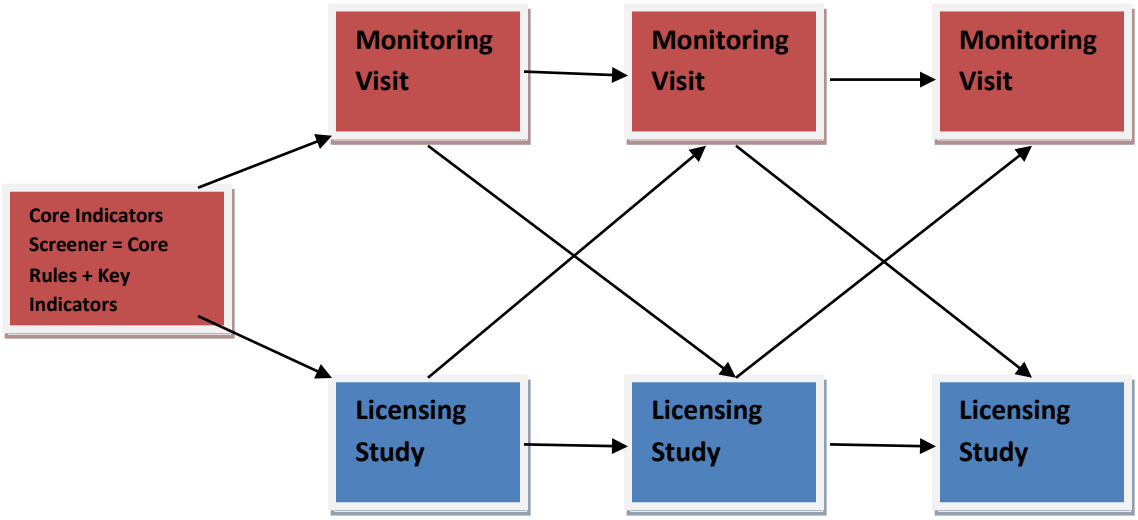
A confounding of data occurred with the first two validation approaches because the Core Rules were influenced a great deal by the National Child Care Key Indicators (NCCKI) (Fiene, 2002) where 10 of the 13 Core Rules overlapped significantly with the NCCKI. This helped to increase the correlation between the Core Rules and the Licensing Studies because the Core Rules represented both risk assessment and key indicator rules. Using both risk assessment and key indicator rules together is an ideal differential monitoring approach (Fiene, 2012). Most states use one or the other but generally not together. By including the newly generated key indicators from this study where there is also overlap with the NCCKI, it should enhance the differential monitoring approach utilized by DECAL.

[2] ACDW decisions were compared with using severity as a factor and not using it as a factor in the scoring system with Core Rules. No significant differences were found between the two scoring systems; therefore, the results in this study represent Core Rule scores without severity included since this is the simpler model.

RECOMMENDATIONS

The following recommendations[3] can be made from this Licensing Differential Monitoring Validation Study.

1) **First research question/validation recommendation:** Revise the worksheet determination scoring relative to the visiting protocol by combining the Core Rules with a Key Indicator approach so that if any of the Core Rules or Key Indicators are out of compliance, then a full compliance review (Licensing Study) should be used. The present worksheet determination scoring protocol is overly complex. Just moving to a more comprehensive review (Licensing Study) based on non-compliance with the Core Rules will simplify the scoring protocol and make determinations more straightforward. If there is full (100%) compliance with the Core Rules and Key Indicators, then the next scheduled review of the program would be an abbreviated Monitoring Visit. If there is not 100% compliance with the Core Rules and Key Indicators, then the next scheduled review of the program would be a Licensing Study reviewing all child care rules. Based upon the compliance/non-compliance scores of the Licensing Study will determine how often the program will be visited. A revised Georgia Differential Monitoring System could potentially look like the following:



Compliance Decisions:

Core Indicators *= Core Rules + Key Indicators – this becomes a screening tool to determine if a program receives a Licensing Study reviewing all child care rules or an abbreviated Monitoring visit continuing to review key indicator and core rules for their next visit.*

Core Indicators (100%) *= the next visit is a Monitoring Visit.. Every 3-4 years a full Licensing Study is conducted.*

Core Indicators (not 100%) *= The next visit is a Licensing Study where all rules are reviewed.*

Compliance *= 96%+ with all rules and 100% with Core Indicators. The next visit is a Monitoring Visit.*

Non-compliance *= less than 96% with all rules. The next visit is a Licensing Study..*

2) **Second research question/validation recommendation:** Follow the development of weighted risk assessment tools as outlined by Fiene & Kroh (2000) in the *NARA Licensing Chapter* for CCC and FCC. It has been over 20 years since Core Rules were weighted. It is recommended that Core Rules be weighted every 10 years. Doing a weighted risk assessment would help confirm that the present Core Rules are the highest risk rules.

3) **Third research question/validation recommendation:** Confirm the CCC (ERS/PS) and FCC results by conducting a more recent program quality study that reflects all the changes made within the CCC and FCC systems. Although FCC program quality and Licensing Study and Core Rules reached statistical significance, the overall correlation was too low (Licensing Studies = .19; Core Rules = .17). With the CCC system the Pre-K program demonstrated significant correlations between ERS/PK and Licensing Study (.48) & Core Rules (.60) but not the Preschool program (ERS/PS: Licensing Studies = .21; Core Rules = .27).

[3] These recommendations are drawn from the data in this study and previous studies conducted by the author in which the empirical evidence led to similar recommendations.

REFERENCES

Fiene (2012). *Differential monitoring logic model and algorithm (DMLMA©).* Middletown, PA: Research Institute for Key Indicators.

Fiene (2007). Child Development Program Evaluation & Caregiver Observation Scale, in T Halle (Ed.), *Early Care and Education Quality Measures Compendium*, Washington, D.C.: Child Trends.

Fiene (2003). Licensing related indicators of quality child care, *Child Care Bulletin*, Winter 2002-2003, pps 12-13.

Fiene (2002). *Thirteen indicators of quality child care: Research update*. Washington, DC: Office of the Assistant Secretary for Planning and Evaluation, US Department of Health and Human Services.

Fiene (1985). Measuring the effectiveness of regulations, *New England Journal of Human Services, 5*(2), 38-39.

Fiene (1981). A new tool for day care monitoring introduced by children's consortium, *Evaluation Practice, 1*(2), 10-11.

Fiene & Kroh (2000). Licensing Measurement and Systems, *NARA Licensing Curriculum*. Washington, D.C.: National Association for Regulatory Administration.

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

Griffin & Fiene (1995). *A systematic approach to policy planning and quality improvement for child care: A technical manual for state administrators*. Washington, D.C.: National Center for Clinical Infant Programs-Zero to Three.

Harms, T., Cryer, D.R., & Clifford, R.M. (2007). Family child care environment rating scale: Revised edition. New York: Teachers College Press.

Harms, T., Cryer, D., & Clifford, R. M. (2006). *Infant/toddler environment rating scale: Revised edition*. New York: Teachers College Press.

Harms, T., Clifford, R. M., & Cryer, D. (2005). *Early childhood environment rating scale: Revised edition*. New York: Teachers College Press.

Maxwell, K. L., Early, D. M., Bryant, D., Kraus, S., Hume, K., & Crawford, G. (2009a). Georgia study of early care and education: Child care center findings. Chapel Hill, NC: The University of North Carolina at Chapel Hill, FPG Child Development Institute.

Maxwell, K. L., Early, D. M., Bryant, D., Kraus, S., Hume, K., & Crawford, G. (2009b). *Georgia study of early care and education: Findings from Georgia's Pre-K Program.* Chapel Hill, NC: The University of North Carolina at Chapel Hill, FPG Child Development Institute.

Maxwell, K. L., Early, D. M., Bryant, D., Kraus, S., & Hume, K., (2010). Georgia study of early care and education: Family child care findings. Chapel Hill: The University of North Carolina at Chapel Hill, FPG Child Development Institute.

Morgan, Stevenson, Fiene, & Stephens (1986). Gaps and excesses in the regulation of child day care, *Reviews of Infectious Diseases--Infectious Diseases in Child Day Care: Management and Prevention, 8*(4), 634-643.

Zellman, G. L. and Fiene, R. (2012). *Validation of Quality Rating and Improvement Systems for Early Care and Education and School-Age Care,* Research-to-Policy, Research-to-Practice Brief OPRE 2012. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

TABLE 8 - FOUR APPROACHES TO VALIDATING A QRIS (Zellman & Fiene, 2012)

Approach	Activities and Purpose	Typical Questions Approach Addresses	Issues and Limitations
1. Examine the validity of key underlying concepts	Assess whether basic QRIS quality components and standards are the "right" ones by examining levels of empirical and expert support.	Do the quality components capture the key elements of quality? Is there sufficient empirical and expert support for including each standard?	Different QRISs may use different decision rules about what standards to include in the system.
2. Examine the measurement strategy and the psychometric properties of the measures used to assess quality	Examine whether the process used to document and verify each indicator is yielding accurate results. Examine properties of key quality measures, e.g., inter-rater reliability on observational measures, scoring of documentation, and inter-item correlations to determine if measures are psychometrically sound. Examine the relationships among the component measures to assess whether they are functioning as expected. Examine cut scores and combining rules to determine the most appropriate ways to combine measures of quality standards into summary ratings.	What is the reliability and accuracy of indicators assessed through program administrator self-report or by document review? What is the reliability and accuracy of indicators assessed through observation? Do quality measures perform as expected? (e.g., do subscales emerge as intended by the authors of the measures?) Do measures of similar standards relate more closely to each other than to other measures? Do measures relate to each other in ways consistent with theory? Do different cut scores produce better rating distributions (e.g., programs across all levels rather than programs at only one or two levels) or more meaningful distinctions among programs?	This validation activity is especially important given that some component measures were likely developed in low-stakes settings and have not been examined in the context of QRIS.

TABLE 8 (CONTINUED)

Approach	Activities and Purpose	Typical Questions Approach Addresses	Issues and Limitations
3. Assess the outputs of the rating process	Examine variation and patterns of program-level ratings within and across program types to ensure that the ratings are functioning as intended. Examine relationship of program-level ratings to other quality indicators to determine if ratings are assessing quality in expected ways. Examine alternate cut points and rules to determine how well the ratings distinguish different levels of quality.	Do programs with different program-level ratings differ in meaningful ways on alternative quality measures? Do rating distributions vary by program type, e.g., ratings of center-based programs compared to ratings of home-based programs? Are current cut scores and combining rules producing appropriate distributions across rating levels?	These validation activities depend on a reasonable level of confidence about the quality components, standards and indicators as well as the process used to designate ratings.
4. Examine how ratings are associated with children's outcomes.	Examine the relationship between program-level ratings and selected child outcomes to determine whether higher program ratings are associated with better child outcomes.	Do children who attend higher-rated programs have greater gains in skills than children who attend lower-quality programs?	Appropriate demographic and program level control variables must be included in analyses to account for selection factors. Studies could be done on child and program samples to save resources. Findings do not permit attribution of causality about QRIS participation but inferences can be made about how quality influences children's outcomes.

FIGURE 1- DIFFERENTIAL MONITORING LOGIC MODEL AND ALGORITHM (Fiene, 2012)
***DMLMA*© Applied to the Georgia Child Care Licensing System**

CI + PQ => RA + KI => DM

Georgia Examples:

CI = Comprehensive Tool = Licensing Study (LS – All Rules)
PQ = Program Quality Tool = Environmental Rating Scales (ERS = ECERS-R, ITERS-R, FCCERS-R)
RA = Risk Assessment Tool = Core Rules (CR)
KI = Key Indicators Tool = presently Georgia does not have a KI
DM = Differential Monitoring Tool = ACDW (Compliance/Non-Compliance Decision)

A very important concept in this validation study is that the system employed by DECAL is a risk assessment approach rather than a key indicator methodology which is based upon predictor rules. The *DMLMA*© is a new methodology assessing the effectiveness and efficiency of Differential Monitoring systems being used by state regulatory agencies and provides the conceptual model for this study.

DMLMA© Thresholds:
High Correlations (.70+) = CI x KI.
Moderate Correlations (.50+) = CI x RA; RA x DM; RA x KI; KI x DM.
Lower Correlations (.30+) = PQ x CI; PQ x RA; PQ x KI.

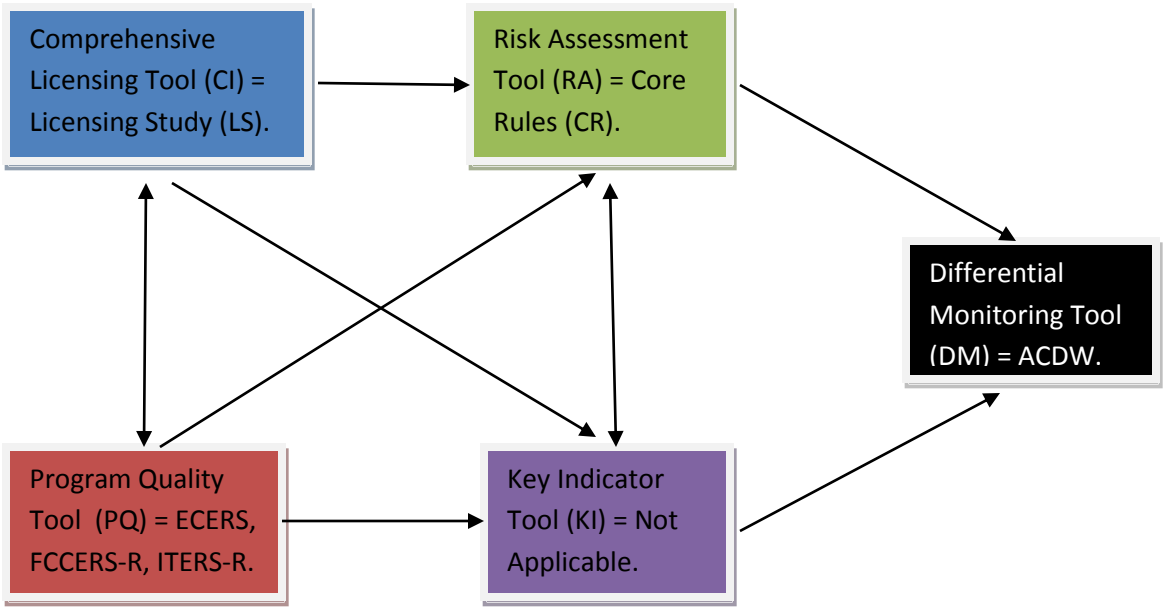


Table 9 - Listing of Key Indicators for Georgia Child Care Centers with Phi Coefficients

591-1-1-25 (3) requires that the center and surrounding premises be clean, free of debris and in good repair. (Phi = .49)

591-1-1-.25 (13) requires that hazardous equipment, materials and supplies be inaccessible to children. (Phi = .46)

591-1-1-.26 (6) requires that outdoor equipment be free of hazards such as lead-based paint, sharp corners, rust and splinters. (Phi = .44)

591-1-1-.26 (8) requires the playground to be kept clean, free of litter and hazards. (Phi = .59)

591-1-1.26 (7) requires that a resilient surface be provided and maintained beneath the fall zone of climbing and swinging equipment. (Phi = .57)

591-1-1-.36 (6)(a-c) requires the center to maintain on the vehicle current information for each child including a) center and passenger information; b) emergency medical information and c) a passenger checklist. (Phi = .49)

591-1-1-.14 (1) requires that at least 50% of the caregiver staff have current first aid and CPR training. (Phi = .49)

591-1-1-.08 (a)-(f) requires the center to maintain a file for each child while such child is in care and for one year after that child is no longer enrolled…. (Phi = .44)

Table 10 - Listing of Key Indicators for Georgia Family Child Care Homes with Phi Coefficients

290.2.3-.11(2)(C) requires that fire drills be practiced monthly and shall be documented and kept on file for one year. (Phi = .51)

290-2-3-.11 (2)(f) requires that poisons, medicines, cleaning agents and other hazardous materials be in locked areas or inaccessible to children. (Phi = .61)

290-2-3-.11 (1)(f) requires the family day care home and any vehicle used to have a first aid kit….. (Phi = .57)

290-2-3-.07 (4) requires that the provider obtain ten clock hours of training in child care issues from an approved source within the first year and thereafter on an annual basis. (Phi = .58)

290-2-3-.08 (1)(a) requires the family day care home to maintain a file for each child that includes the child's name, birth date, parents or guardian's name, home and business addresses and telephone numbers. (Phi = .63)

290-2-3-.08 (1)(b) requires that the record for each child contain the names(s), address(es) and telephone number(s) of person(s) to contact in emergencies when the parent cannot be reached. (Phi = .57)

290-2-3-.08 (1)(b) requires the family day care home to maintain a file for each child that includes the name, address and telephone number of the child's physician to contact in emergencies. (Phi = .55)

290-2-3-.08 (1)(f) requires the family day care home to maintain a file for each child that includes known allergies, physical problems, mental health disorders, mental retardation or developmental disabilities which would limit the child's participation in the program. (Phi = .51)

290-2-3-.08 (1)(c) requires the family day care home to maintain a file for each child that includes evidence of age appropriate immunizations or a signed affidavit against such immunizations; enrollment in the home may not continue for more than 30 days without such evidence. (Phi = .72)

Table 11 - Key Indicator Formula Matrix for Generating Key Indicators*

	Providers In Compliance on Rule	*Programs Out Of Compliance on Rule*	*Row Total*
High Group **	***A***	***B***	***Y***
Low Group* **	**C**	**D**	***Z***
Column Total	***W***	***X***	**Grand Total**

(* This computation occurred for each licensing rule)

****

Figure 2 - Key Indicator Statistical Methodology (Calculating the Phi Coefficient)

$$\phi = (A)(D) - (B)(C) \div \sqrt{(W)(X)(Y)(Z)}$$

A = High Group + Programs in Compliance on Specific Rule.
B = High Group + Programs out of Compliance on Specific Rule.
C = Low Group + Programs in Compliance on Specific Rule.
D = Low Group + Programs out of Compliance on Specific Rule.
W = Total Number of Programs in Compliance on Specific Rule.
X = Total Number of Programs out of Compliance on Specific Rule.
Y = Total Number of Programs in High Group.
Z = Total Number of Programs in Low Group .
****High Group = Top 25% of Programs in Compliance with all Rules.***
*****Low Group = Bottom 25% of Programs in Compliance with all Rules.***

****

Table 12 – Phi Coefficient Decision Table

Phi Coefficient Range	**Characteristic of Indicator**	**Decision**
(+1.00) – (+.26)	**Good Predictor**	**Include**
(+.25) – (-.25)	**Unpredictable**	**Do not Include**
(-.26) – (-1.00)	**Terrible Predictor**	**Do not Include**

Table 13 - Comparison of the Pre-K and Preschool Programs

Compliance Level*	Pre-K ECERS-R**(N)	Preschool ECERS-R***(N)
100	4.88 (4)	3.40 (15)
99	4.13 (6)	4.35 (7)
98	4.38 (6)	3.89 (13)
97	3.99 (4)	3.15 (9)
96	4.36 (2)	3.16 (13)
95	4.60 (2)	3.53 (5)
90	3.43 (2)	2.56 (5)
80	2.56 (1)	2.38 (2)

***Compliance Level = the number of child care rule violations subtracted from 100.**

100 = Full Compliance with Rules
99-98 = Substantial Compliance with Rules
97-90 = Medium Level of Compliance with Rules
80 = Low Level of Compliance with Rules

****Pre-K ECERS-R = average score of Pre-K Program classrooms as compared to the respective compliance levels. (N) = Sample Size.**

*****Preschool ECERS-R = average score of Preschool Program classrooms as compared to the respective compliance levels. (N) = Sample Size.**

From this comparison there is more of a linear relationship between compliance levels and ECERS-R average scores for Pre-K Program classrooms than with the Preschool Program classrooms where there is more of a curvilinear or plateau effect at the upper end of compliance levels (Full Compliance). In order to attain the necessary correlational thresholds (+.30+) for validation for the third approach to validation, having a linear relationship rather than curvilinear will enhance this occurring. When a curvilinear or plateau effect occurs there is too great a likelihood that programs at a medium level of quality will be introduced into the highest (full) level of compliance. From a public policy standpoint this is an undesirable result.

The other item to note with the data distributions is that the Preschool ECERS-R data are more restricted than the Pre-K Program ECERS-R data. In other words, there is less variance in the Preschool Program ECERS-R data than in the Pre-K Program ECERS-R data.

There is an important limitation in these data that the reader must be aware of in not drawing any conclusions that the presence of a Pre-K Program classroom in any way is causing the change in licensing compliance. There is a relationship between the two but there is no assumption of causality.

Georgia Licensing Validation Technical Elements Appendix

Because of the nature of this report being a state's first attempt at fully validating it's Child Care Licensing Core Rule Differential Monitoring Approach utilizing the Zellman & Fiene (2012) Validation Framework and Fiene's DMLMA (2012) Model, certain questions surfaced regarding the terminology and the methodology being used in this report. This Technical Elements Appendix provides answers to specific questions that have been raised regarding these methodologies.

1. *How were the multiple years of data handled?*

The Licensing Study data used to make the comparisons are the facility reports that were the earliest facility observations so that these data would be closest to when the program quality data were collected. The other more recent Licensing Studies were not used in this comparison.

2. *If the Core Rules, Key Indicator, and Licensing Study values are counts of violations, how was the fact that different sites had different numbers of visits handled?*

Because only the earliest Licensing Study data was used, the number of visits were not an issue in the scoring.

3. *If the Core Rules, Key Indicator, and Licensing Study values are counts of violations, were all levels of violation risk (low, medium, high, extreme) handled the same?*

Yes, there were very few occurrences of high and extreme in the data base and also no significant differences were found when a sample of the rule violations with and without the levels of violation risk were compared. Therefore the simpler formula in which levels of violation risk were not used was selected.

4. *How did you determine the minimum correlations (DMLMA thresholds) for each analysis? Was this computed separately for this analysis or are the minimum correlations based on previous work?*

The DMLMA thresholds were determined from previous research work conducted by the author of this study on this model over the past 30 years. These were the average correlational thresholds that have been proposed for making validation determinations. The reason for utilizing the DMLMA model and thresholds is that the Zellman & Fiene (2012) Framework provides guidance in how to select specific validation approaches, what are the specific questions answered by the approach and what are the limitations of the particular approach. The DMLMA model builds upon this but provides a suggested scoring protocol by comparing correlational thresholds in a specific state to historical trends.

5. *Was Phi calculated for every rule in the licensing study? Can the full list be added to the appendix?*

Yes, Phi was calculated for every rule in the licensing study but most of them could not be computed because there was so few rule violations in the majority of the rules. This is typical of state licensing data sets and the full Phi comparisons are not depicted because it does not add any information to the state report.

6. How did you determine which of the Licensing Study rules should be counted as Key Indicators?

The Key Indicator statistical methodology based upon a specific cut off point for the Phi Coefficient in which the p values were .0001 or less. This is a very stringent cut off point but it has been found historically that the p values needed to be lowered as the data distributions became more skewed with programs overall compliance levels increasing over time.

7. How were sites that had no infant/toddler (i.e., no ITERS score) handled for the third validation approach? How were sites that had only a GA Pre-K (no preschool) handled?

For scoring purposes only those facilities that had both the ECERS and ITERS scores were used in making comparisons with the licensing data related to the third approach to validation. The GA Pre-K were scored and compared in the same way.

8. On Table 13, why is the number of violation subtracted from 100 (rather than from the maximum possible)?

Generally this scoring is done because it is more intuitive to think in terms of 100% in compliance as a score of "100" rather than a score of "0". This conversion is used in all state licensing reports that involve the DMLMA, Key Indicators and Risk Assessment Models.

FOR ADDITIONAL INFORMATION REGARDING THIS REPORT AND STUDY:

Bentley Ponder, Ph.D.
DECAL Director of Research and Evaluation
Bentley.Ponder@decal.ga.gov

Richard Fiene, Ph.D.
Research Institute for Key Indicators *LLC*
DrFiene@gmail.com

OFFICE OF HEAD START KEY INDICATOR PROJECT REPORT

Richard Fiene, Ph.D.

The purpose of this report is to present to the Office of Head Start (OHS) Key Indicators of their Head Start Performance Standards (HSPS) that have the ability to statistically predict substantial compliance with all Compliance Measures and ultimately the majority of HSPS's. The analytical and methodological basis of this approach is based upon a *Differential Monitoring Logic Model and Algorithm (DMLMA©)* (Fiene, 2012) (see Appendix 3). The DMLMA© is the 4th generation of an Early Childhood Program Quality Indicator Model (ECPQIM)(Fiene & Nixon, 1985; Griffin & Fiene, 1995; Fiene & Kroh, 2000). Only a portion of the *DMLMA©* model was utilized in this report which focused on key indicators, risk assessment, and program quality.

Definitions:

Risk Assessment (RA) - a differential monitoring approach that employs using only those rules, standards, or regulations that place children at greatest risk of mortality or morbidity if violations/citations occur with the specific rule, standard, or regulation.

Key Indicators (KI) - a differential monitoring approach that employs using only those rules, standards, or regulations that statistically predict overall compliance with all the rules, standards, or regulations. In other words, if a program is 100% in compliance with the Key Indicators the program will also be in substantial to full compliance with all rules, standards, or regulations. The reverse is also true in that if a program is not 100% in compliance with the Key Indicators the program will also have other areas of non-compliance with all the rules, standards, or regulations.

Differential Monitoring (DM) - this is a relatively new approach to determining the number of visits made to programs and what rules, standards, or regulations are reviewed during these visits. There are two measurement tools that drive differential monitoring, one is Weighted Risk Assessment tools and the other is Key Indicator checklists. Weighted Risk Assessments determine how often a program will be visited while Key Indicator checklists determine what rules, standards, or regulations will be reviewed in the program. Differential monitoring is a very powerful approach when Risk Assessment is combined with Key Indicators because a program is reviewed by the most critical rules, standards, or regulations and the most predictive rules, standards, or regulations. See Appendix 3 which presents a Logic Model & Algorithm for Differential Monitoring *(DMLMA©)(*Fiene, 2012).

Program Quality (PQ) - for the purposes of this study this was measured via the CLASS – Classroom Assessment Scoring System. The CLASS has three sub-scales (ES = Emotional Support, CO = Classroom Organization, and IS = Instructional Support). The CLASS is a tool that is identified in the research literature as measuring classroom quality similar to the ERS tools.

Early Childhood Program Quality Indicator Model (ECPQIM) – these are models that employ a key indicator or dashboard approach to program monitoring. Major program monitoring systems in early care and education are integrated conceptually so that the overall early care and education system can be assessed and validated. With these models, it is possible to compare results obtained from licensing systems, quality rating and improvement systems (QRIS), risk assessment systems, key indicator systems, technical assistance, and child development/early learning outcome systems. The various approaches to validation are interposed within this model and the specific expected correlational thresholds that should be observed amongst the key elements of the model are suggested. **Key Elements** of the model are the following (see Appendix 3 for details): **CI** = state or federal standards, usually rules or regulations that measure health and safety - *Caring for Our Children or Head Start Performance* Standards will be applicable here. **PQ** = Quality Rating and Improvement Systems (QRIS) standards at the state level; *ERS (ECERS, ITERS, FDCRS), CLASS, or CDPES* (Fiene & Nixon, 1985). **RA** = risk assessment tools/systems in which only the most critical rules/standards are measured. *Stepping Stones* is an example of this approach**. KI** = key indicators in which only predictor rules/standards are measured. The *Thirteen Indicators of Quality Child Care* is an example of this approach. **DM** = differential monitoring decision making in which it is determined if a program is in compliance or not and the number of visits/the number of rules/standards are ascertained from a scoring protocol. **PD** = technical assistance/training and/or professional development system which provides targeted assistance to the program based upon the **DM** results. **CO** = child outcomes which assesses how well the children are developing which is the ultimate goal of the system.

The organization of this report is as follows:

1) **The first section** will provide an overall analysis the Head Start (HS), Early Head Start (EHS), and Head Start/Early Head Start (HS/EHS) programs[1,4] ;

2) **The second section** will provide analyses of the various content areas (CA) within the HSPS[4];

3) **The third section** will provide analyses of the relationship between the HSPS as measured by compliance with the Compliance Measures (CM) and the program quality scores (CLASS scores)[3];

4) **The fourth and final section** will provide the analyses that produced the key indicators (KI) and recommendations in how it could be used.[2]

The source of data for this report is all the Tri-Annual On-Site Monitoring visits for 2012 which consisted of 422 reviews of programs across the country. There were 191 Head Start (HS) only programs, 33 Early Head Start (EHS) only programs, and 198 Head Start/Early Head Start (HS/EHS) programs reviewed. This is a representative sample of Head Start and Early Head Start programs nationally representing approximately 25% of the total number of Head Start programs.

Before proceeding with the results of this study, a few clarifying and definitional terms need to be highlighted. In the *2012 edition of OHS On-Site Review Protocol* and the *2013 OHS Monitoring Protocol*, Compliance Indicators (CI) and Key Indicators (KI) are respectively mentioned. In the licensing literature, when the term "Indicators" is used it refers to standards/rules that are predictive of overall compliance with all rules/standards. However, as defined by OHS, indicators (CI/KI) are used within the context of risk assessment which means that these indicators are the standards which are most important/critical

to the OHS in their monitoring reviews. These indicators therefore are not predictive in essence. That is the focus of this report/study which is to determine which of these indicators are predictive of overall compliance with all the compliance/key indicators. This is a common misconception in the human service regulatory field where risk assessment tools and key indicator tools purposes are confused. As we move forward please keep the definitions in mind related to the distinctions and functionality of risk assessment and key indicators.

For the purposes of this study, 131 Compliance Measures (CM), organized into seven (7) Content Areas (CA), were reviewed and analyzed. The seven content areas are the following: Program Governance; Management Systems; Fiscal Integrity; Eligibility, Recruitment, Selection, Enrollment, and Attendance; Child Health and Safety; Family and Community Engagement; Child Development and Education. Ten CM's were from Program Governance (GOV), 10 were from Management Systems (SYS), 22 were from Fiscal Integrity (FIS), 11 were from Eligibility, Recruitment, Selection, Enrollment, and Attendance (ERSEA), 34 were from Child Health and Safety (CHS), 16 were from Family and Community Engagement (FCE), and 28 were from Child Development and Education (CDE)[4].

Section 1 - Head Start (HS), Early Head Start (EHS), and Head Start/Early Head Start (HS/EHS) programs

In order to determine if analyses needed to be performed separately on Head Start (HS), Early Head Start (EHS), and Head Start/Early Head Start (HS/EHS) combined programs, the first series of analyses were performed to determine if any statistically significant differences existed amongst these three groups. This is a very important first analysis because it will help to determine the stability of the sample selected and of the overall system. In other words, is there a good deal of consistency across all service types: HS, EHS, and HS/EHS.

Based upon Table 1, no statistically significant differences were determined amongst the three groups (HS, EHS, HS/EHS) with Compliance Measures (CM) or CLASS (ES, CO, IS) Scores indicating that using the full 422 sample and not having to do separate analyses for the three groups was the correct analytical framework. However, where it is appropriate, any statistically significant differences amongst the various program types will be highlighted.

Table 1 – Head Start, Early Head Start, & Head Start/Early Head Start With CM and CLASS/ES, CO, IS

Program Type	CM(N)	CLASS/ES(N)	CLASS/CO(N)	CLASS/IS(N)
Head Start (HS)	3.72(191)	5.88(186)	5.43(186)	2.97(186)
Early Head Start (EHS)	2.67(33)	-----*	-----*	-----*
Head Start (HS/EHS)	3.07(198)	5.91(198)	5.47(198)	3.00(198)
Totals	**3.33(422)**	**5.89(384)**	**5.45(384)**	**2.98(384)**
Statistical Significance	**NS**	**NS**	**NS**	**NS**

CM = Compliance Measures (Average Number of Violations)
CLASS/ES = CLASS Emotional Support Average Score
CLASS/CO = CLASS Classroom Organization Average Score
CLASS/IS = CLASS Instructional Support Average Score
NS = Not Significant
N = Number of Programs
****CLASS data were not collected in EHS.***

The average number of violations with the Compliance Measures for Head Start (3.72), Early Head Start (2.67) and Head Start/EHS (3.07) was not significant in utilizing a One-Way ANOVA. There were 191 Head Start (HS) programs, 33 Early Head Start (EHS) programs, and 198 Head Start (HS/EHS) programs.

Comparisons were also made with Head Start and Head Start/EHS on the various CLASS sub-scales (ES = Emotional Support, CO = Classroom Organization, and IS = Instructional Support) and no significant differences were found between these two groups. The EHS (n = 33) was not used because CLASS data were not collected in these programs.

The practical implication of the above results is that the same monitoring tools and the resulting Head Start Key Indicator (HSKI) to be developed as a result of this study can be used in the three main types of programs: Head Start, Early Head Start, and Head Start/EHS. There is no need to have separate tools.

Section 2 - Content Areas

The second series of analyses was to look more closely at the 7 content areas (CA) to measure demographically any differences amongst the various areas. In order to do this a weighted average had to be determined in order to compare the various areas because of the differences in the number of Compliance Measures (CM) used in each content area. Table 2 provides the results of these analyses. For the total sample of 422 sites, Management Systems (SYS) Content Area (CA) had the highest number of violations with the Compliance Measures (CM) with 359. The SYS/CA also had the highest average number of violations with 35.90 because there were only 10 CM. For the total sample of 422 sites, the lowest number of violations was in the Family and Community Engagement (FCE) Content Area (CA) with 48 violations with CM. It also had the lowest average number of violations with 3.00.

For the Head Start only sites (n = 191), a similar distribution as with the total sample (n = 422) is depicted in which Management Systems (SYS) Content Area (CA) had the highest number of violations with the Compliance Measures (CM) with 192. The SYS/CA also had the highest average number of violations with 19.20 because again there were only 10 CM. The lowest number of violations was in the Family and Community Engagement (FCE) Content Area (CA) with 20 violations with CM. It also had the lowest average number of violations with 1.25.

For the Early Head Start only (n = 33) and the Head Start/Early Head Start (n = 198) sites, the ranking of the various Content Areas changed somewhat with the total number of violations and the average number of violations from the Total Sample (n = 422) and the Head Start only (n = 191) sites but not dramatically. For example, the Family and Community Engagement (FCE); Child Development and Education (CDE); and the Eligibility, Recruitment, Selection, Enrollment, and Attendance (ERSEA) Content Areas switched rankings in which it had the fewest total violations and the average number of violations (see Table 2).

Table 2 – Comparing Content Areas and Program Types

	Total Violations/(Rank)				Average # of Violations/(Rank)				CM
Content Areas	**TOT**	**HS**	**EHS**	**HS/EHS**	**TOT**	**HS**	**EHS**	**HS/EHS**	**#**
FCE	48(1)	20(1)	2(1)	26(2)	3.00(1)	1.25(1)	0.125(1)	1.63(2)	16
ERSEA	62(2)	37(2)	6(3)	19(1)	5.64(3)	3.36(3)	0.545(3)	1.73(3)	11
CDE	91(3)	43(3)	5(2)	43(3)	3.25(2)	1.54(2)	0.179(2)	1.54(1)	28
GOV	150(4)	94(4)	6(3)	50(4)	15.00(6)	9.40(6)	0.600(4)	5.00(5)	10
FIS	255(5)	114(5)	23(7)	118(5)	11.59(5)	5.18(5)	1.045(6)	5.36(6)	22
CHS	333(6)	151(6)	22(6)	160(7)	9.79(4)	4.44(4)	0.647(5)	4.71(4)	34
SYS	359(7)	192(7)	20(5)	147(6)	35.90(7)	19.20(7)	2.000(7)	14.70(7)	10

CONTENT AREAS (CA):
FCE = FAMILY and COMMUNITY ENGAGEMENT
ERSEA = ELIGIBILITY, RECRUITMENT, SELECTION, ENROLLMENT, and ATTENDANCE
CDE = CHILD DEVELOPMENT AND EDUCATION
GOV = PROGRAM GOVERNANCE
FIS = FISCAL INTEGRITY
CHS =CHILD HEALTH AND SAFETY
SYS = MANAGEMENT SYSTEMS

TOT = TOTAL NUMBER OF SITES, FULL SAMPLE OF 422 SITES
HS = HEAD START ONLY PROGRAMS
EHS = EARLY HEAD START ONLY PROGRAM
HS/EHS = HEAD START AND EARLY HEAD START COMBINED PROGRAMS
CM = NUMBER OF COMPLIANCE MEASURES

TOTAL VIOLATIONS = ALL THE VIOLATIONS FOR A SPECIFIC CONTENT AREA.
AVERAGE # OF VIOLATIONS = THE TOTAL VIOLATIONS FOR A SPECIFIC CA DIVIDED BY THE NUMBER OF COMPLIANCE MEASURES FOR THAT SPECIFIC CONTENT AREA.
RANK = HOW EACH CONTENT AREA COMPARES TO THE OTHER CONTENT AREAS FOR THE RESPECTIVE PROGRAM TYPE.

For the total sample (n = 422), other CA's had different configurations between the total number of violations and the average number of violations as demonstrated by CHS – Child Health and Safety in which there was a total of 333 violations but the average number of violations was 9.79 because there were 34 Compliance Measures (CM). Program Governance (GOV) had 150 total violations and a weighted-average of 15 violations with 10 CM. Child Development and Education (CDE) had 91 total violations and a weighted-average of 3.25 violations. Fiscal Integrity (FIS) had 255 total violations and a weighted-average of 11.59 violations. And lastly, Eligibility, Recruitment, Selection, Enrollment, and Attendance (ERSEA) had 62 total violations and a weighted-average of 5.64 violations.

The Head Start only (HS = 191), Early Head Start only (EHS = 33), and the Head Start/Early Head Start (HS/EHS = 198) programs followed a similar pattern as with the total sample (n = 422). This indicates a great deal of consistency in the sample drawn. See Appendix 4 for violation data for all 131 Compliance Measures.

The practical implication of the above findings is that certain Content Areas (SYS, GOV, FIS) may need additional exploration by OHS because of their high rates of non-compliance with the Compliance Measures.

Section 3 – Program Quality

This section provides comparisons between the Compliance Measures (CM) data and the CLASS (ES, CO, IS) data. This is a very important section because there is always the concern that compliance with the HSPS has no relationship to program quality as measured by the CLASS. In Table 3, correlations were run between the CM data and the CLASS scores for Emotional Support (ES), Classroom Organization (CO), and Instruction Support (IS) for the Head Start only and the Head Start/Early Head Start programs. The EHS only programs were not included because CLASS data are not collected on these programs. The results are very positive and statistically significant in most cases. It is also important to note the very positive correlation between the Head Start Key Indicators (HSKI[2]) and CLASS. This result supports using the HSKI in monitoring Head Start.

Table 3 – Relationship Between Compliance Measures (CM), KI, and CLASS (ES, CO, IS) Scores

	Compliance Measures Content Areas								Key Indicators
CLASS	**CM**	***FCE***	***ERSEA***	***CDE***	***GOV***	***FIS***	***CHS***	***SYS***	**KI**
CLASS/ES	.22**	.13*	.15**	.15**	.11*	.05	.23**	.17**	.27**
CLASS/CO	.19**	.13*	.11*	.16**	.04	.06	.21**	.15**	.25**
CLASS/IS	.20**	.10	.12*	.12*	.13*	.06	.18**	.11*	.17**

CM Violations = Total Compliance Measure Violations

CONTENT AREAS (CA):
FCE = FAMILY and COMMUNITY ENGAGEMENT
ERSEA = ELIGIBILITY, RECRUITMENT, SELECTION, ENROLLMENT, and ATTENDANCE
CDE = CHILD DEVELOPMENT AND EDUCATION
GOV = PROGRAM GOVERNANCE
FIS = FISCAL INTEGRITY
CHS =CHILD HEALTH AND SAFETY
SYS = MANAGEMENT SYSTEMS

CLASS/IS = Average CLASS IS (Instructional Support) Score
CLASS/ES = Average CLASS ES (Emotional Support) Score
CLASS/CO = Average CLASS CO (Classroom Organization) Score

KI = Key Indicators Total Score

*** p < .01***
*** p < .05***

See Appendix 6 & 6A for the inter-correlations amongst all the Content Areas, HSKI, and Total Compliance with Compliance Measures.

These results are very important but it is equally important to look more specifically at the distribution of the Compliance Measures (CM) scores and their relationship to the CLASS data (see Appendix 5 for detailed graphic distributions and Appendix 6 & 6A for the inter-correlations amongst all the CA). When this is done a very interesting trend appears (see Table 3a) in which a definite plateau occurs as the scores move from more violations or lower compliance with the Compliance Measures (25-20 to 3-8 CM Violations) to fewer violations or substantial compliance with the Compliance Measures (1-2 CM Violations) and full compliance with the Compliance Measures (Zero (0) CM Violations).

Table 3a – Aggregate Scores Comparing CM Violations with CLASS Scores

CM Violations		IS	ES	CO	Number/Percent
0	(Full Compliance)	3.03	5.99	5.59	75/19%
1-2	(Substantial Compliance)	3.15	5.93	5.50	135/35%
3-8	(Mid-Compliance)	2.87	5.85	5.37	143/40%
9-19	(Lower Compliance)	2.65	5.71	5.32	28/6%
20-25	(Lowest Compliance)	2.56	5.52	4.93	3/1%
Significance		**F = 4.92; p < .001**	**F = 4.918; p < .001**	**F = 4.174; p < .003**	

CM Violations = Compliance Measure Violations (lower score = higher compliance)(higher score = lower compliance)
IS = Average CLASS IS (Instructional Support) Score
ES = Average CLASS ES (Emotional Support) Score
CO = Average CLASS CO (Classroom Organization) Score
#/% = Number of programs and Percent of programs at each level of compliance

When comparing these groupings in Table 3a the results from a One Way ANOVA were significant ($F = 4.92$; $p < .001$) for the CLASS/IS Scores. The average CLASS/IS Score when there were no CM Violations was 3.03. The average CLASS/IS Score when there were 1-2 CM Violations was 3.15. The average CLASS/IS Score when there were 3-8 CM Violations was 2.87. The average CLASS/IS Score when there were 9-19 CM Violations was 2.65. And finally, the average CLASS/IS Score when there were 20-25 violations was 2.56. The results were very similar with the CLASS/ES and CLASS/CO scores as well in which the results from a One Way ANOVA were statistically significant for the CLASS/ES ($F = 4.918$; $p < .001$) and for the CLASS/CO ($F = 4.174$; $p < .003$). These results clearly demonstrate that being in full or substantial compliance with the Compliance Measures correlates with more positive scores on the CLASS. Approximately 55% of the Head Start programs are at the full or substantial compliance level.

The practical implication of the above findings is that placing equal emphasis on full as well as substantial compliance with the Compliance Measures could be an acceptable public policy decision.

Section 4 – Head Start Key Indicators (HSKI)

The fourth and final section of this report is in some ways the most important since this is the focus of the study: developing statistically predictive Key Indicator (KI) Compliance Measures (CM) – the Head Start Key Indicators (HSKI).

These are the statistically predictive Key Indicators based upon the KI methodology, correlations with the CLASS/ES, CO, IS, and correlations with the CM Total Violation scores. Table 4 lists the results while Appendix 1 has the specific KI's content specified. Appendix 2 depicts the KI Formula Matrix. Only those Compliance Measures (CM) that had significant results on three of the five correlations were selected to be Head Start Key Indicator Compliance Measures (HSKI).

The methodology used to generate the Compliance Measure Key Indicators sorted the top 20% of programs in compliance and compared this group to the bottom 27% of programs in compliance. The middle 53% of programs were not used in order to determine the Key Indicators. These cut off points

were determined by the compliance distribution in which 20% of the programs were in 100% compliance while 27% of the programs had compliance scores of 95% or less.

Table 4 – Head Start Key Indicator (HSKI) Compliance Measures (CM) and CLASS and Total Violations

HSKI/CM (2013)	Phi	CLASS/ES	CLASS/CO	CLASS/IS	Total Violations
CDE4.1	.28***	.10*	ns	ns	.30***
CHS1.1	.39***	.15**	.16**	ns	.39***
CHS1.2	.33***	.18**	.15**	.10*	.36***
CHS2.1	.49***	.18**	.15**	ns	.54***
CHS3.10	.39***	.11*	.11*	ns	.24***
GOV2.1	.31***	.11*	ns	ns	.46***
SYS2.1	.47***	.15**	.16**	.14**	.55***
SYS3.4	.58***	.13*	.10*	ns	.36***

Phi = the phi coefficient which statistically predicts compliance with the full set of CM's.

CLASS/ES = correlations between the specific CM and this specific scale of the CLASS.
CLASS/CO = correlations between the specific CM and this specific scale of the CLASS.
CLASS/IS = correlations between the specific CM and this specific scale of the CLASS.

Total Violations = correlations between the specific CM and the total number of CM violations for each program.

** p < .05*
*** p < .01*
**** p < .001*
ns = not significant

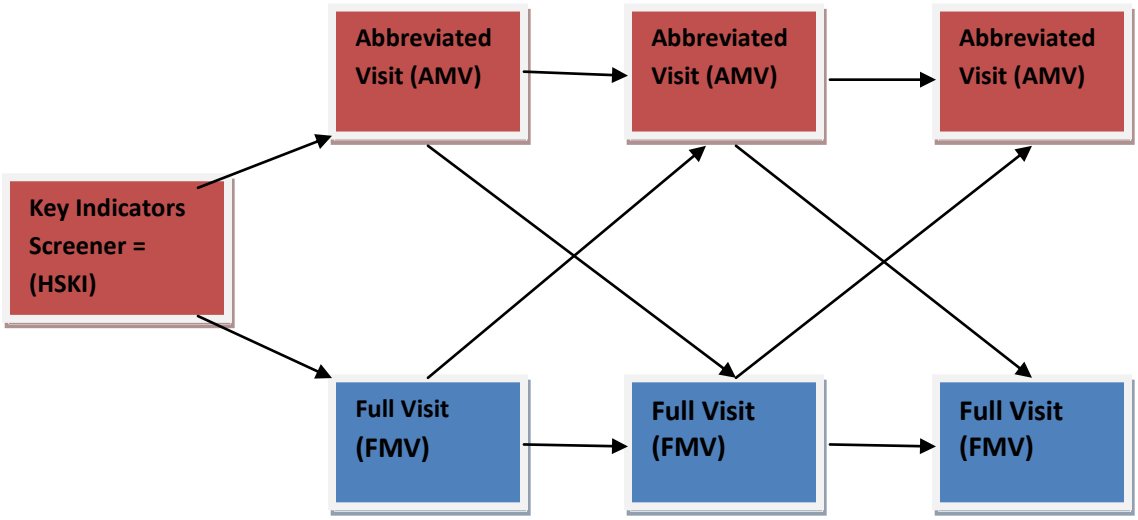
Separate Key Indicators were run for just Head Start only and Head Start/Early Head Start programs but the key indicators were only a subset of the above list, albeit a shorter list in each case. Based upon those phi coefficients, it was determined that using the above list for all Head Start only, Early Head Start, and Head Start/Early Head Start was a more efficient and effective way to monitor all the programs with one list of indicators rather than having separate key indicators for program types. The separate phi coefficients run for Head Start only and Head Start/Early Head Start programs did not show any significant differences because they were sub-samples of the overall sample drawn.

Section 4A – Suggested Use of the HSKI for Head Start Program Monitoring

Now that Key Indicators have been generated, the next question is how to use HSKI in the program monitoring of Head Start. A possible way in which the HSKI could be used would be the following (see Figure 1) in which a differential monitoring approach could be used:

All programs would be administered the HSKI. If there is full (100%) compliance with the Head Start Key Indicators (HSKI) then the next scheduled review of the program would be an Abbreviated Monitoring Visit (AMV). If there is not 100% compliance with the Head Start Key Indicators (HSKI) then the next scheduled review of the program would be a Full Monitoring Visit (FMV) in which all Compliance Measures are reviewed. Based upon the results of the FMV a determination could be made regarding a compliance or non-compliance decision (see Figure 1) and how often the program will be visited.

Figure 1 – Head Start Key Indicator (HSKI) Compliance Measures Differential Monitoring Model



Compliance Decisions:

Head Start Key Indicators (HSKI) – this becomes a screening tool to determine if a program receives an AMV OR FMV visit.

***HSKI (100%)** = For the next visit, an Abbreviated Monitoring Visit (AMV) is conducted. Every 3-4 yrs a full Monitoring is conducted.*

***HSKI (not 100%)** = For the next visit, a Full Monitoring Visit (FMV) is conducted and all CMs are reviewed.*

***Compliance** = 98%+ with all CMs which indicates substantial to full compliance and 100% with HSKI. For the next visit, an Abbreviated Monitoring Visit (AMV) is conducted.*

***Non-compliance** = less than 98% with all CMs which indicates low compliance. For the next visit a Full Monitoring Visit (FMV) is conducted.*

Moving to a differential monitoring system could provide a cost effective and efficient model for Head Start program monitoring. This revision to the Head Start program monitoring system would combine a risk assessment and key indicator approach (see Appendix 3) in determining what compliance measures to review, how often, and how comprehensive a review should be utilized. It would continue to focus on the most critical compliance measures that statistically predict overall compliance with the full complement of compliance measures.

See Appendix 7 – Figure 2 for how the above differential monitoring system could impact the present Head Start Tri-Annual Review Monitoring System. In this appendix, a cost neutral monitoring system is proposed based upon the above DMLMA/Key Indicator Model.

References

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

Fiene (2003). Licensing related indicators of quality child care, *Child Care Bulletin*, Winter 2002-2003, pps 12-13.

Fiene (1985). Measuring the effectiveness of regulations, *New England Journal of Human Services, 5*(2), 38-39.

Fiene (1981). A new tool for day care monitoring introduced by children's consortium, *Evaluation Practice, 1*(2), 10-11.

Fiene & Kroh (2000). Licensing Measurement and Systems, *NARA Licensing Curriculum*. Washington, D.C.: National Association for Regulatory Administration.

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

Griffin & Fiene (1995). *A systematic approach to policy planning and quality improvement for child care: A technical manual for state administrators*. Washington, D.C.: National Center for Clinical Infant Programs-Zero to Three.

Footnotes

1) PIR Dashboard Key Indicators could not be generated because the PIR data demonstrated little statistical predictive ability to be useful for discriminating between high and low compliant programs or program quality with the exception of staff having CDA's.

2) The correlation between Compliance Measures (CM) and the statistically predictive Key Indicators (HSKI) was .77 which exceeds the expected correlation threshold.

3) The correlations between the CLASS/ES, CO, IS and Key Indicators were the following: .27, .25, .17 respectively. The correlations between KI and ES and CO were higher than the correlations between CM and ES, CO as reported earlier in this report. The correlation between IS and CM was higher .20 than KI and IS (.17).

4) Because this study spans the 2012 Review Protocol and 2013 Monitoring Protocol, Compliance Indicators and Compliance Measures are used interchangeably with a preference given to using Compliance Measures (CM) in this report. There are 139 Compliance Indicators; 115 Compliance Measures, but for the purposes of this study 131 Compliance Measures were available in the 2012 Head Start data base drawn for this study.

For additional information regarding this report, please contact:

Richard Fiene, Ph.D., Director
Research Institute for Key Indicators
41 Grandview Drive
Middletown, PA. 17057
DrFiene@gmail.com
RIKI.Institute@gmail.com
717-944-5868 Phone and Fax
http://RIKInstitute.wikispaces.com
http://pennstate.academia.edu/RickFiene

February 2013 (revised March, April 2013)

Appendix 1 – Head Start Key Indicators (HSKI) Compliance Measures Content

CM	Content	Regulations/Law
CDE4.1*	The program hires teachers who have the required qualifications, training, and experience.	1304.52(f), 645A(h)(1), 648A(a)(3)(B)(i), 648A(a)(3)(B)(ii), 648A(a)(3)(B)(iii)
CHS1.1	The program engages parents in obtaining from a health care professional a determination of whether each child is up to date on a schedule of primary and preventive health care (including dental) and assists parents in bringing their children up to date when necessary and keeping their children up to date as required.	1304.20(a)(1)(ii), 1304.20(a)(1)(ii)(A), 1304.20(a)(1)(ii)(B)
CHS1.2	The program ensures that each child with a known, observable, or suspected health, oral health, or developmental problem receives follow-up and further testing, examination, and treatment from a licensed or certified health care professional.	1304.20(a)(1)(iii), 1304.20(a)(1)(iv), 1304.20(c)(3)(ii)
CHS2.1	The program, in collaboration with each child's parent, performs or obtains the required linguistically and age-appropriate screenings to identify concerns regarding children within 45 calendar days of entry into the program, obtains guidance on how to use the screening results, and uses multiple sources of information to make appropriate referrals.	1304.20(a)(2), 1304.20(b)(1), 1304.20(b)(2), 1304.20(b)(3)
CHS3.10	Maintenance, repair, safety of facility and equipment	1304.53(a)(7)
GOV2.1*	Members of the governing body and the Policy Council receive appropriate training and technical assistance to ensure that members understand information they receive and can provide effective oversight of, make appropriate decisions for, and participate in programs of the Head Start agency.	642(d)(3)
SYS2.1	The program established and regularly implements a process of ongoing monitoring of its operations and services, including delegate agencies, in order to ensure compliance with Federal regulations, adherence to its own program procedures, and progress towards the goals developed through its Self-Assessment process.	1304.51(i)(2), 641A(g)(3)
SYS3.4	Prior to employing an individual, the program obtains a: Federal, State, or Tribal criminal record check covering all jurisdictions where the program provides Head Start services to children; Federal, State, or Tribal criminal record check as required by the law of the jurisdiction where the program provides Head Start services; Criminal record check as otherwise required by Federal law	648A(g)(3)(A), 648A(g)(3)(B), 648A(g)(3)(C)

**** FY 2013 Office of Head Start Monitoring Protocol (October 26, 2013) Compliance Measures***

Appendix 2: Key Indicator Formula Matrix for HSKI – Head Start Key Indicators

	Providers In Compliance	*Programs Out Of Compliance*	*Row Total*
High Group	***A***	***B***	***Y***
Low Group	**C**	**D**	***Z***
Column Total	***W***	***X***	**Grand Total**

Key Indicator Statistical Methodology (Calculating the Phi Coefficient):

$$\phi = (A)(D) - (B)(C) \div \sqrt{(W)(X)(Y)(Z)}$$

A = High Group + Programs in Compliance on Specific Compliance Measure.
B = High Group + Programs out of Compliance on Specific Compliance Measure.
C = Low Group + Programs in Compliance on Specific Compliance Measure.
D = Low Group + Programs out of Compliance on Specific Compliance Measure.

W = Total Number of Programs in Compliance on Specific Compliance Measure.
X = Total Number of Programs out of Compliance on Specific Compliance Measure.
Y = Total Number of Programs in High Group.
Z = Total Number of Programs in Low Group.

High Group = Top 20% of Programs in Compliance with all Compliance Measures.
Low Group = Bottom 27% of Programs in Compliance with all Compliance Measures.

Phi Coefficient Range	Characteristic of Indicator	Decision
(+1.00) – (+.26)	**Good Predictor**	**Include on HSKI**
(+.25) – (0)	**Too Easy**	**Do not Include**
(0) – (-.25)	**Too Difficult**	**Do not Include**
(-.26) – (-1.00)	**Terrible Predictor**	**Do not Include**

Appendix 3

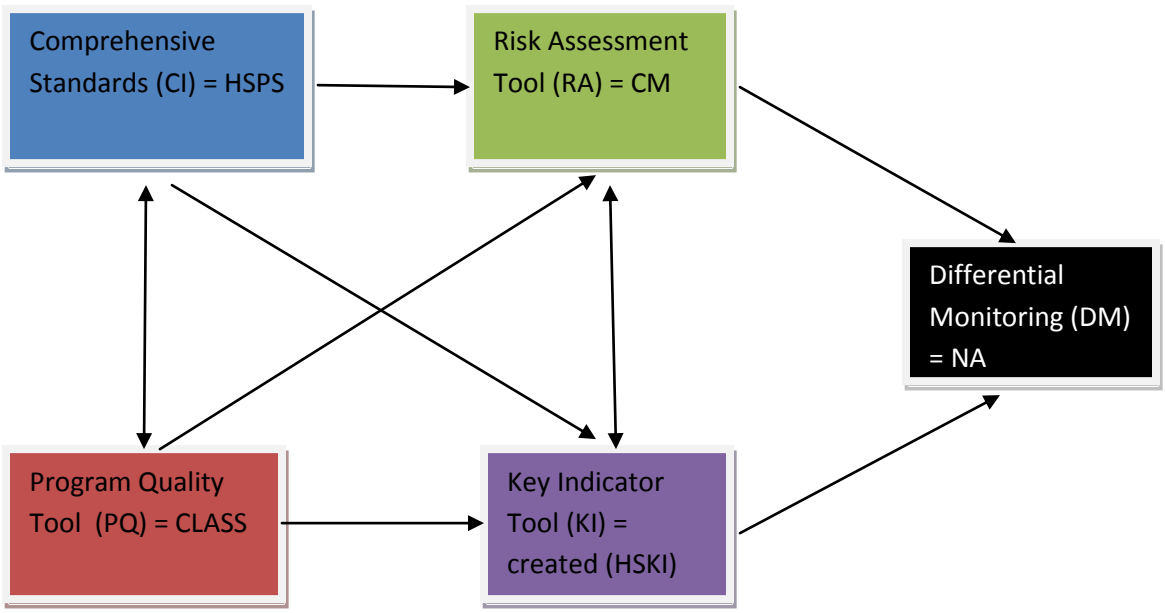
DIFFERENTIAL MONITORING LOGIC MODEL AND ALGORITHM (Fiene, 2012) *DMLMA©* Applied to the Office of Head Start Program Monitoring Compliance System

CI + PQ => RA + KI => DM

Head Start Examples:

CI = Head Start Performance Standards (HSPS)
PQ = CLASS ES, IS, CO (CLASS)
RA = Compliance Measures (CM)
KI = Key Indicators (generated from this study = Head Start Key Indicators (HSKI))
DM = Not Applicable at this time (NA) but see Figure 1 for a proposed model

DMLMA© Thresholds:
High Correlations (.70+) = CI x KI.
Moderate Correlations (.50+) = CI x RA; RA x DM; RA x KI; KI x DM.
Lower Correlations (.30+) = PQ x CI; PQ x RA; PQ x KI.



Appendix 4: Content Areas and Compliance Measures

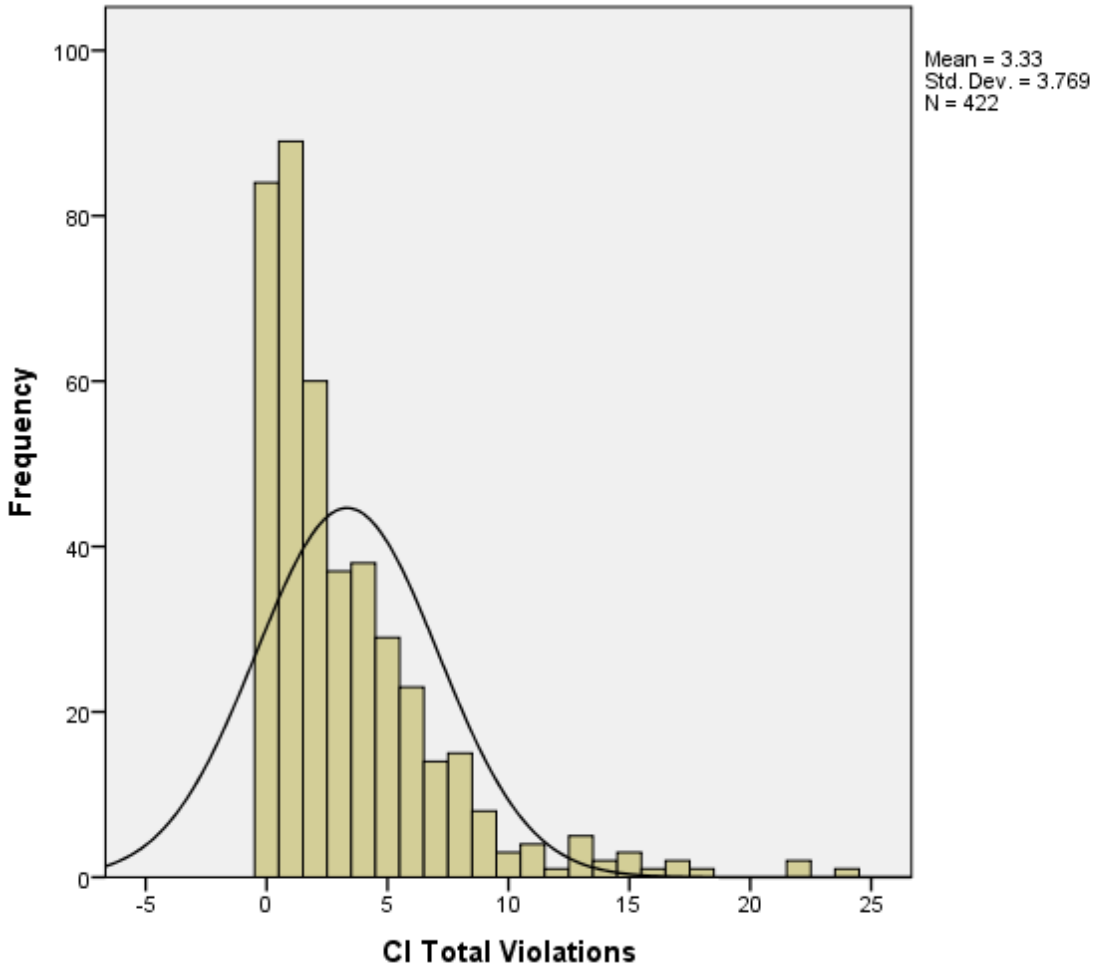
Content Areas and Compliance Measures ***FY 2012 OHS On-Site Review Protocol (FY 2013 OHS Monitoring Protocol)***	**Percent (%) Compliance**
CDE - CHILD DEVELOPMENT AND EDUCATION	**99%**
1.1(2.2) The program implements a curriculum that is aligned with the Head Start Child Development and Early Learning Framework…	99%
1.2 The program implements a curriculum that is evidence-based…	99%
1.3(2.1) The curriculum is comprehensive….	99%
2.1 The program implements an infant toddler curriculum….	99%
2.2 The program develops secure relationships in out of home care settings for infants and toddlers…	100%
2.3 The program implements an infant/toddler curriculum that encourages trust….	100%
2.4 The program encourages the development of self-awareness, autonomy…..	100%
2.5 The program fosters independence.	100%
2.6 The program enhances each child's strengths by encouraging self control….	99%
2.7 The program plans for routines and transitions…..	99%
2.9 The program encourages respect for others feelings and rights.	99%
2.10 The program provides opportunities for children to engage in child-initiated…..	100%
2.11 Nutrition services contribute to children's development and socialization…..	100%
3.1 The program uses information from screenings, ongoing observations…..	99%
3.3 The programs' nutrition program is designed and implemented to meet the nutritional needs….	98%
3.4(CHS4.5) Meal and snack periods are appropriately scheduled….	99%
3.5(3.2) Services provided to children with identified disabilities are designed to support…..	100%
3.6(3.3) The program designates a staff member or consultant to coordinate services for children w/disabilities…	100%
3.7(3.4) The program has secured the services of a mental health professional…..	97%
3.8(3.5) The program's approach to CDE is developmentally and linguistically appropriate….	99%
4.1 The program establishes goals for improving school readiness…..	98%
4.2 The program uses self assessment information on school readiness goals…..	99%
4.3 The program demonstrates that children who are dual language learners…..	100%
5.1(4.1) The program hires teachers who have the required qualifications, training, & experience.	92%
5.2 The program ensures that family child care providers have the required qualifications….	100%
5.3 The program ensures that all full time Head Start employees who provide direct education….	96%
5.4 The program ensures that home visitors have the required qualifications, training….	99%
5.5 When the majority of children speak the same language…..	99%
CHS - CHILD HEALTH AND SAFETY	**97%**
1.1 The program engages parents in obtaining from a health care professional a determination of whether each child….	89%
1.2 The program ensures that each child with a known, observable, or suspected health, oral health…..	92%
1.3 The program involves parents, consulting with them immediately when child health or developmental problems…..	100%
1.4 The program informs parents and obtains authorization prior to all health procedures….	98%
1.5 The program has established procedures for tracking the provision of health services.	97%
1.6 The EHS program helps pregnant women, immediately after enrollment in the program, access through referrals…..	100%
1.7 Program health staff conduct a home visit or ensure that a health staff member visits each newborn within 2 weeks of birth….	97%
2.1 The program, in collaboration with each child's parent, performs or obtains the required screenings….	84%
2.2 A coordinated screening, assessment, and referral process for all children….	98%
2.3 The program, in partnership with the LEA or Part C Agency, works to inform and engage parents in all plans for screenings….	99%
3.1 Facilities used for center based program options comply with state and local licensing….	100%
3.2 The program ensures that sufficient equipment, toys, materials, and furniture are provided….	97%
3.3 Precautions are taken to ensure the safety of children.	99%
3.4 The program ensures that medication is properly stored and is not accessible to children.	98%
3.5 The program ensures that no hazards are present around children.	89%
3.6 The program ensures that sleeping arrangements for infants do not use soft bedding materials.	99%
3.7 All infant and toddler toys are made of non-toxic materials and sanitized regularly.	99%
3.8 The program has adequate usable indoor and outdoor space.	99%
3.9 Outdoor play areas are arranged to prevent children from getting into unsafe or unsupervised areas…..	100%
3.10 The program provides for maintenance, repair, safety, and security of all Head Start facilities and equipment.	85%
3.11 The program's facilities provide adequately for children with disabilities…..	100%
4.1 Staff, volunteers, and children wash their hands with soap and running water.	98%
4.2 Spilled bodily fluids are cleaned up and disinfected immediately….	100%
4.3 The program adopts sanitation and hygiene practices for diapering……	99%

4.4(4.7) The program ensures that facilities are available for proper refrigerated storage and handling of breast milk and formula.	100%
4.5(4.8) Effective oral hygiene is promoted among children in conjunction with meals.	99%
5.1 The program ensures appropriate class and group sizes based on the predominant age of the children.	99%
5.2 The program ensures that no more than eight children are placed in an infant and toddler space…..	99%
6.1 The program's vehicles are properly equipped.	99%
6.2 At least one bus monitor is aboard the vehicle at all times.	99%
6.3 Children are released only to a parent……	99%
6.4 Each bus monitor, before duty, has been trained on child boarding and exiting procedures……	99%
6.5 The program ensures that persons employed to drive vehicles receive the required behind the wheel training….	99%
6.6 Specific types of transportation assistance offered are made clear to all prospective families…	100%
ERSEA – ELIGIBILITY, RECRUITMENT, SLECTION, ENROLLMENT, AND ATTENDANCE	**98%**
1.1 The program developed and implemented a process that is designed to actively recruit families…..	99%
1.2 The program has a systematic process for establishing selection criteria……	99%
1.3 The program has established and implemented outreach and enrollment policies and procedures….	99%
2.1 Program staff verified each child's eligibility……	94%
2.2 The program enrolls children who are categorically eligible…..	99%
2.3 The American Indian or Alaskan Native programs ensure that the children who meet the following requirements….	100%
3.1 Actual program enrollment is composed of at least 10 percent children with disabilities.	96%
3.2 The program enrolled 100% of its funded enrollment…..	98%
3.3 The program has documentation to support monthly enrollment data …..	98%
4.1 When monthly average daily attendance in center based programs falls below 85%, the causes of absenteeism….	99%
4.2 The program ensures that no child's enrollment or participation in the Head Start program is contingent on payment of a fee.	99%
FCE – FAMILY AND COMMUNITY ENGAGEMENT	**99%**
1.1(1.2) Program staff are familiar with the backgrounds of families and children…..	100%
1.2(1.3) A strength based and family driven collaborative partnership building process is in place…..	100%
1.3(1.4) The program provides resources and services for families' needs, goals, and interests…..	99%
2.1 The program provides opportunities for parents to enhance their parenting skills…..	99%
2.2 Parents and staff share their respective concerns and observations about their individual children…..	99%
2.3 On site mental health consultation assists the program in providing education to parents……	97%
3.1 Program staff plan, schedule, and facilitate no fewer than two staff parent conferences……	98%
3.2(1.1) The program is open to parents during all program hours….	99%
3.3(3.2) In home based settings, programs encourage parents to be integrally involved in their children's development.	99%
3.4(3.3) Programs provide opportunities for children and families to participate in literacy services……	99%
3.5(3.4) The program builds parents' confidence to advocate for their children by informing parents of their rights…..	99%
4.1 The program has procedures to support successful transitions for enrolled children…..	99%
4.2 The program initiates transition planning for each EHS enrolled child at least 6 months prior to the child's 3rd birthday….	99%
5.1 The program has established and maintains a health services advisory committee.	97%
5.2 The program has taken steps to establish ongoing collaborative relationships with community organizations….	100%
5.3 The program coordinates with and has current interagency agreements in place with LEA's…..	98%
FIS – FISCAL INTEGRITY	**97%**
1.1 The program's financial management systems provide for effective control…..	94%
1.2 The program sought and received prior approval in writing for budget changes….	99%
1.3 The program minimized the time elapsing between the advancement of funds from the Payment Management System….	100%
1.4 The program used Head Start funds to pay the cost of expenses….	99%
1.5 The program has obtained and maintained required insurance coverage for risks and liabilities.	99%
2.1 Financial reports and accounting records are current, accurate, complete….	98%
2.2 Monthly financial statements, are provided to program governing bodies and policy groups….	97%
3.1(3.1) The program has procurement procedures that provide all requirements specified in the applicable statutes…..	95%
3.2(3.1) Contracts and delegate agency agreements are current, available, signed, and dated…..	96%
4.1 Original time records are prepared and properly signed by the individual employee & approved…..	97%
4.2 Head Start or EHS grant funds are not used as any part of the monetary compensation….	99%
4.3 Total compensation for personal services charged to the grant are allowable and reasonable….	98%
5.1 The grantee has implemented procedures to determine allowability, allocability, and reasonableness of costs…..	95%
5.2 Indirect cost charges are supported by a negotiated and approved indirect cost rate.	100%
5.3 If the grantee is required to allocate costs between funding sources, the program utilizes a method for allocating costs….	97%
5.4 The financial records of the grantee are sufficient to allow verification that non-Federal participation is necessary…..	90%
5.5(5.3) The grantee can demonstrate that all contributions of non-Federal share are necessary and reasonable…..	98%
5.6(5.4) During each funding period reviewed the grantee charged to the award only costs resulting from obligations….	98%
6.1(6.1;6.2) For grantees that own facilities purchased or constructed using Head Start grant funds, documentation is available….	97%
6.2(6.1;6.2) The grantee meets property management standards for equipment purchased using HS funds…..	94%
6.3(6.1;6.2) Grantees that entered into a mortgage or other loan agreement using collateral property complied with Federal regs….	97%
6.4(6.1;6.2) The amount which the grantee may claim a cost or non-Federal share contribution……	96%
GOV – PROGRAM GOVERNANCE	**96%**

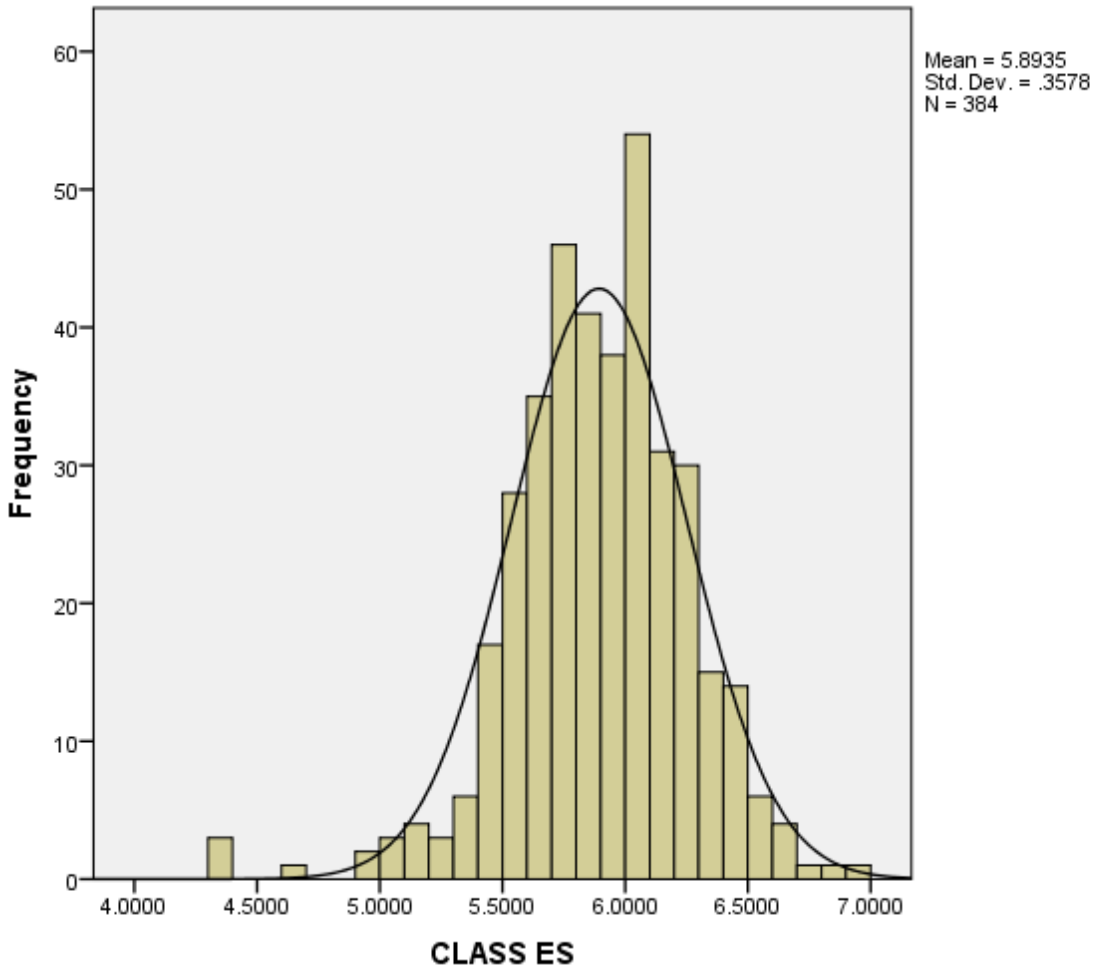
1.1 The program has a governing body….	98%
1.2 The program has established a policy council….	98%
2.1 Policy council and plicy committee members are supported by the program….	99%
2.2 The program has policies and procedures in place to ensure that member of the governing body & PAC are free…..	97%
3.1(2.1) Members of the governing body and the PAC receive appropriate training and TA……	94%
3.2(2.2) The governing body performs required activities and makes decisions pertaining to program administration….	95%
3.3 The governing body approves financial management, accounting, and reporting policies…..	99%
3.4 The governing body reviews and approves all of the program's major policies……	95%
3.5(2.4) The PAC approves and submits decisions about identified program activities to the governing body.	98%
4.1(3.1) Governing body and PAC members r3egulatly receive and use information about program planning…..	88%
SYS – MANAGEMENT SYSTEMS	**91%**
1.1 The program routinely engages in a process of systematic planning that utilizes the results of the community assessment….	97%
1.2(5.1) At least annually, the program conducts a self assessment of program effectiveness….	97%
2.1(5.2) The program established and regularly implements a process of ongoing monitoring of its operations and services….	86%
2.2 The program established and maintains a record keeping system regarding children, families, and staff…..	92%
2.3 The program publishes and makes available to the public an annual report…..	88%
3.1 The program has established an organizational structure that provides for adequate supervision…..	97%
3.2 The program develops and implements written standards of conduct…..	97%
3.3 The program ensures that each staff member completes an initial health examination…..	90%
3.4 Prior to employing an individual, the program obtains: criminal record check….	66%
4.1 The program has mechanisms for regular communication among all program staff….	98%

Appendix 5 – Histograms of Total Compliance Measure Violations, CLASS (IS, ES, CO) Scores and Head Start Key Indicator (HSKI) Scores

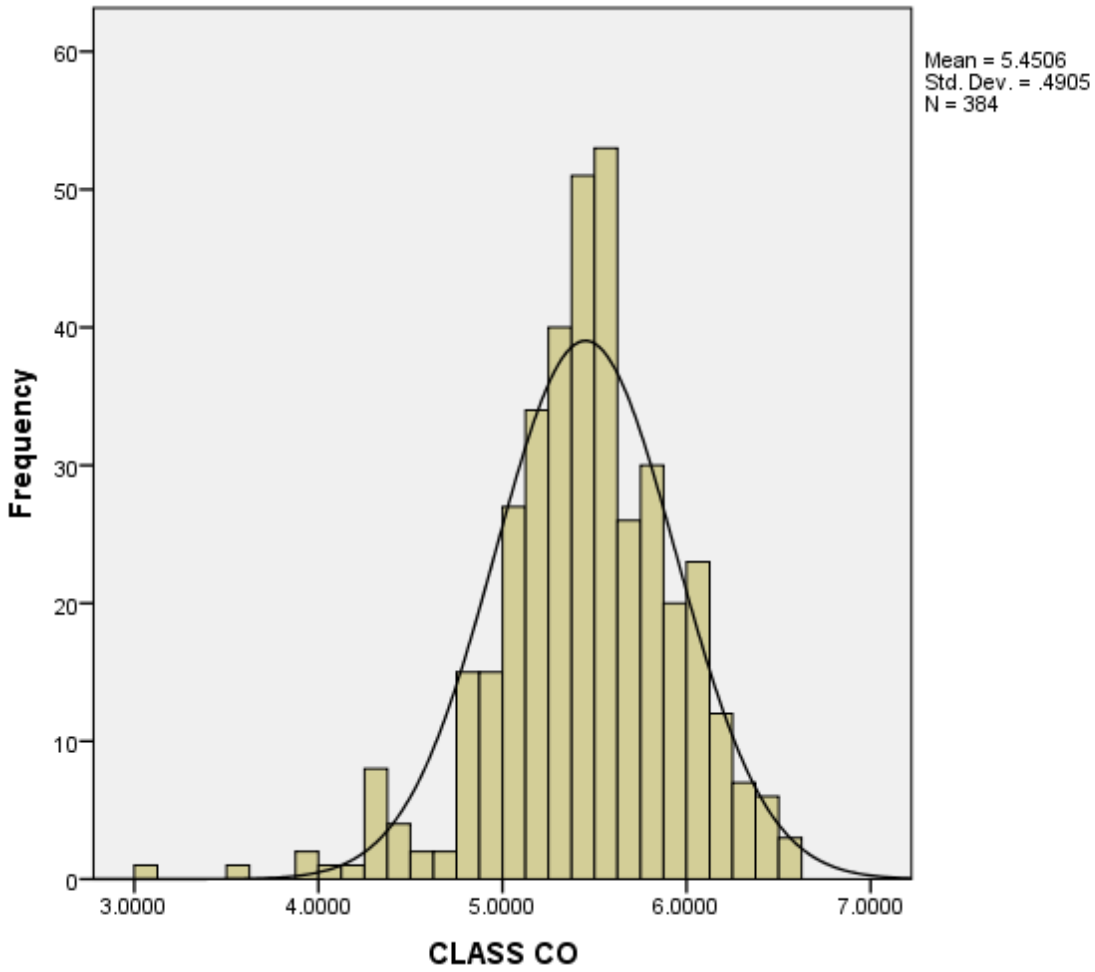
Total Compliance Measure Violations



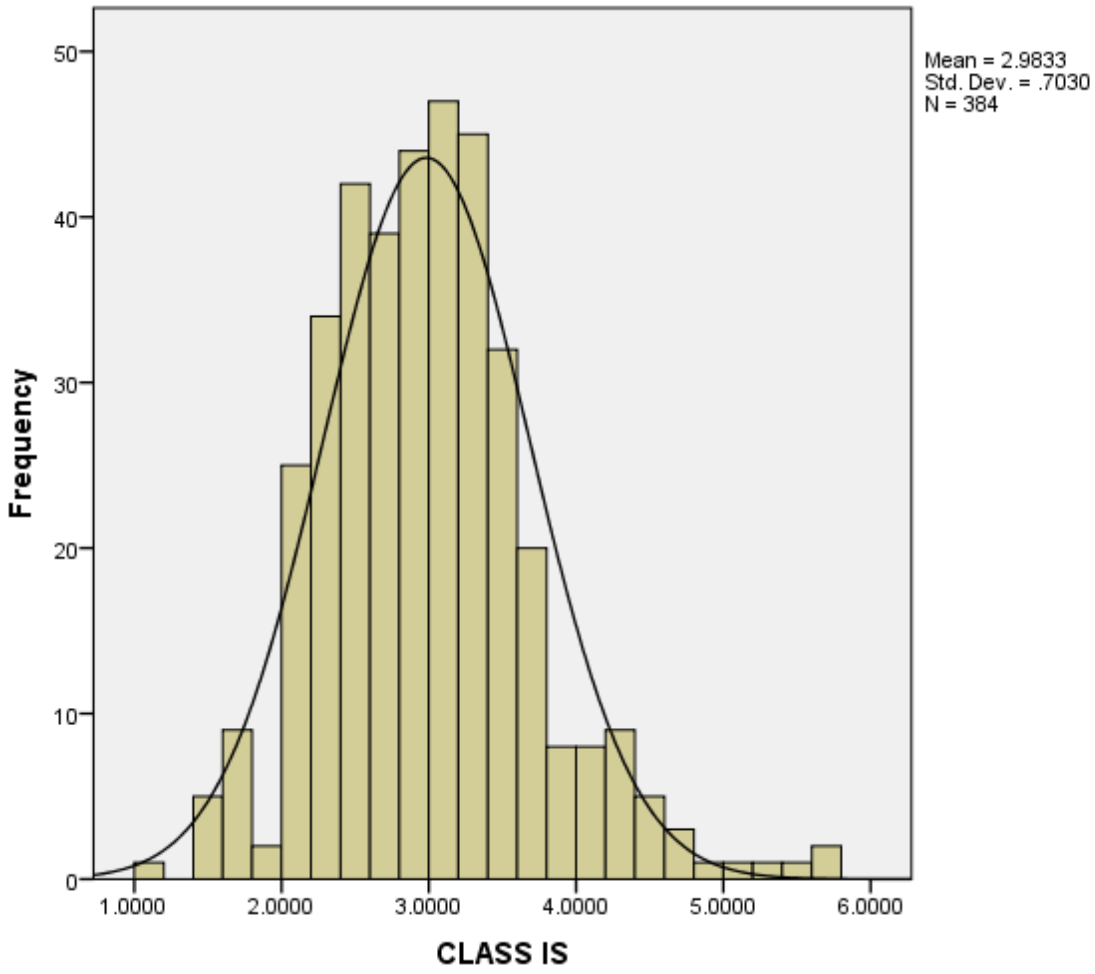
CLASS ES Scores



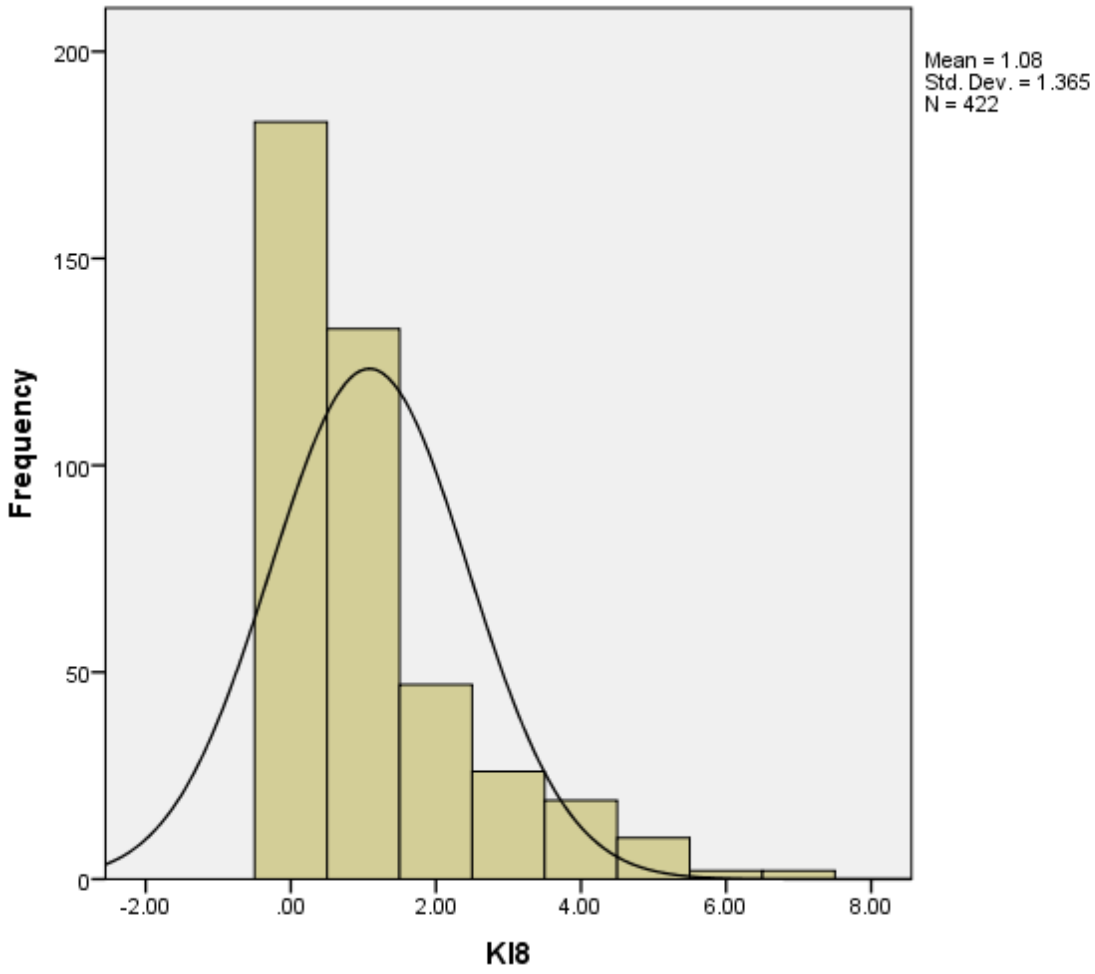
CLASS CO Scores



CLASS IS Scores



Head Start Key Indicators (HSKI) Scores



Appendix 6 -

CONTENT AREA (CA) CORRELATIONS

	CHS	ERSEA	FCE	FIS	GOV	SYS
CDE	.33**	.26**	.06	.14**	.13*	.33**
CHS		.29**	.18**	.09	.25**	.51**
ERSEA			.15**	.10*	.27**	.38**
FCE				.01	.17**	.23**
FIS					.13*	.23**
GOV						.38**

** P < .05*

*** P < .01*

CONTENT AREAS (CA):
FCE = FAMILY and COMMUNITY ENGAGEMENT
ERSEA = ELIGIBILITY, RECRUITMENT, SELECTION, ENROLLMENT, and ATTENDANCE
CDE = CHILD DEVELOPMENT AND EDUCATION
GOV = PROGRAM GOVERNANCE
FIS = FISCAL INTEGRITY
CHS =CHILD HEALTH AND SAFETY
SYS = MANAGEMENT SYSTEMS

Appendix 6A – Total Compliance with Compliance Measures, HSKI, and Content Area Correlations

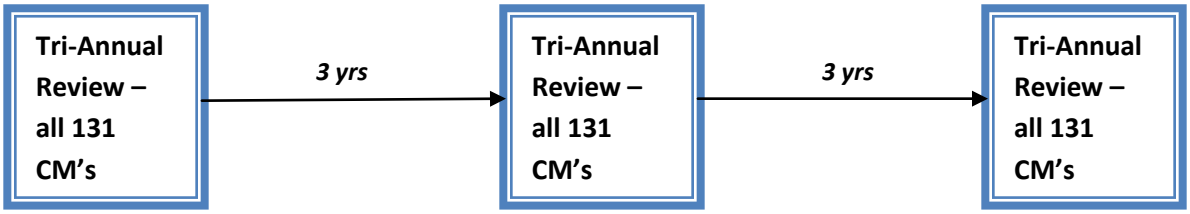
	TOT	HSKI
CDE	.51**	.42**
CHS	.70**	.81**
ERSEA	.49**	.33**
FCE	.30**	.22**
FIS	.50**	.14**
GOV	.57**	.37**
SYS	.78**	.72**

TOT = Total Compliance with all Compliance Measures.
HSKI = Total Compliance with the Head Start Key Indicators.

Appendix 7 – Figure 2 – DMLMA Potential Impact on Tri-Annual Head Start Program Reviews

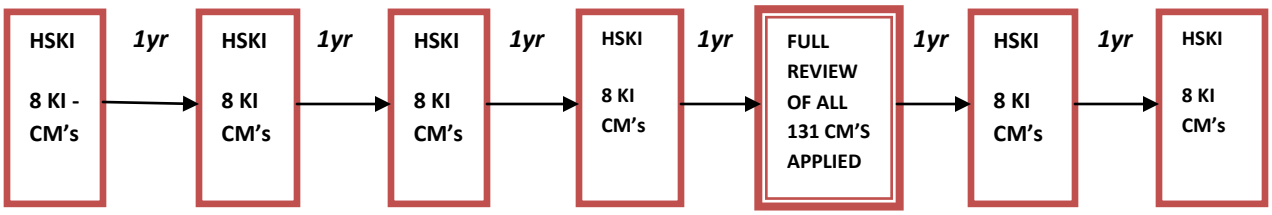
Present Head Start Monitoring System:

All programs receive the same Tri-Annual Reviews regardless of Compliance History:

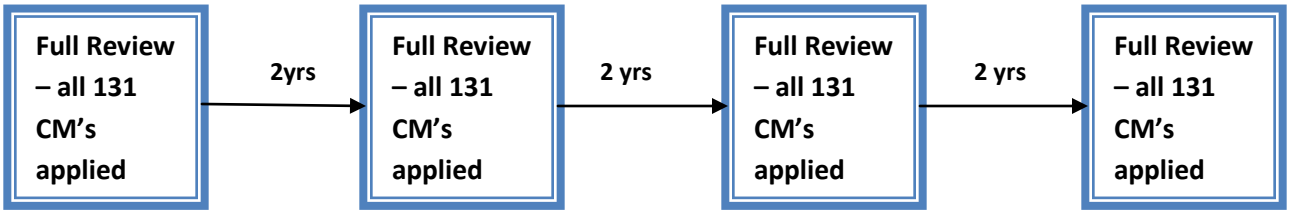


Proposed DMLMA System with Key Indicators (KI):

100% Compliance with the Head Start Key Indicators (HSKI):



If less than 100% with the Head Start Key Indicators (HSKI):



The above proposed change is cost neutral by re-allocating monitoring staff from doing only Tri-Annual Reviews on every program to doing abbreviated monitoring via the HSKI on the highly compliant programs with periodic comprehensive full monitoring less frequently (this would change if a program did not continue to be 100% in-compliance with the HSKI), and only doing more comprehensive full monitoring on those programs with low compliance with the Compliance Measures and/or less than 100% compliance with the HSKI. Once a program was in the high compliance group they would be eligible for the HSKI abbreviated monitoring.

However, the real advantage in this proposed change is the increased frequency of targeted or differential monitoring of all programs.

DMLMA Algorithm with Key Indicators applied to Head Start Tri-Annual Reviews:

Six (6) Years example:

Present Head Start Monitoring System:

(Tri-Annual Visits)(Compliance Measures)(Percent of Programs(%)) = Total Effort
(3)(131)(100) = 39300
Total Effort = ***39300***

Revised Head Start Monitoring DMLMA with Key Indicators System:

100% Compliance with HSKI:
(Number of Monitoring Visits)(Compliance Measures)(Percent of Programs*(%)) = Total Effort
Abbreviated Monitoring Visits using Key Indicators: (6)(8)(43*) = 2064
Full, Comprehensive Monitoring Visit using all Compliance Measures: (1)(131)(43*) = 5633

Less than 100% Compliance with HSKI:
(Number of Monitoring Visits)(Compliance Measures)(Percent of Programs**(%)) = Total Effort
Full, Comprehensive Monitoring Visits using all Compliance Measures: (4)(131)(57**) = 29868

100% Compliance with HSKI + Less than 100% Compliance with HSKI = Total Effort:
Total Effort = 2064 + 5633 + 29868 = ***37565***

---

*This was the actual percent of Head Start Programs that met the criteria of 100% compliance with HSKI in this study.
**This was the actual percent of Head Start Programs that did not meet the criteria of 100% compliance with HSKI in this study.

It would be expected that the total population of Head Start programs would have a similar percent as was found in this representative sample (43% = 100% compliance with HSKI and 57% = less than 100% compliance with HSKI). This representative sample for this study constituted approximately 25% of all Head Start programs nationally.

National Association for Regulatory Administration

Validation of Washington State's Child Care Risk Assessment and Licensing Decision Making Tiered System

Richard Fiene, Ph.D.
Research Psychologist & Senior Research Consultant

June 2020

400 South 4th Street, Ste. 754E | Minneapolis, MN 55415 | (888) 674-7052
www.naralicensing.org
CONSUMER PROTECTION THROUGH PREVENTION

Washington State Department of Children, Youth, and Families Child Care Risk Assessment Licensing Measures and Outputs Validation Study Final Report

Richard Fiene, Ph.D.

National Association for Regulatory Administration

Research Institute for Key Indicators & Penn State University

June 2020

This report will provide the results of two cohorts from a large-scale validation study of Washington State's Department of Children, Youth and Families child care Risk Assessment Licensing Decision Making Tiers System (RALDMTS). The validation involves two key components: 1) Validation of the measurement strategy used to determine the licensing decision making for child care centers and family child care homes; 2) Validation of the licensing system in juxtaposition to the program quality measures (ERS & CLASS) as part of their QRIS – Quality Rating and Improvement System utilized in Washington.

The data set involves two cohorts drawn from licensing reviews in 2019 – 2020. The data reported in this report is from late 2019 and involved 146 sites, and from early 2020 and involved 385 sites. It was driven by the QRIS visiting and assessment schedule.

Let me start by saying that licensing/regulatory compliance data are very different from other data in how they get distributed and therefore should be analyzed. Licensing/regulatory compliance data are grouped into 4 basic buckets: Full regulatory compliance, substantial regulatory compliance, mid-range, and non-optimal regulatory compliance. Obviously full regulatory compliance means 0 violations or 100% compliance with all rules. Substantial regulatory compliance means 1-3 violations with all rules, while low compliance means 10 or move violations with all rules. A middle regulatory compliance range means 4-9 violations with all the rules.

The data were well distributed and fit into the above four (0 - 3) buckets very nicely. Based upon comparing the licensing data to the "Tiers" and "Actions" variables, the licensing decision making system has been validated with high correlations between the licensing data, the Tiers, Risk Assessment Matrix, and the proposed Actions (see Charts 1, 1a and 2, 2a). The data are reported out for both Cohort 1 and then Cohort 2.

With the comparisons between the licensing data and the Environmental Rating Scales (ERS), the licensing data showed the typical "regulatory compliance law of diminishing returns" where the ERS scores were highest with the substantial regulatory compliance range rather than the full regulatory compliance level. In other words, there is not a linear relationship between moving from low to full regulatory compliance and program quality. Programs that are in substantial regulatory compliance and not full regulatory compliance had higher program quality scores. Obviously, the low regulatory compliance programs had also low program quality scores. There is a linear relationship between

regulatory compliance and program quality in moving from low regulatory compliance to the middle and substantial regulatory compliance levels (see Chart 3, 3a).

Chart 1: Tiers By Proposed Actions (Cohort 1)

	Tiers	1	2	3	4
Proposed	**None**	119	0	0	0
Actions	**Tech Assist**	0	12	0	0
	Safety Plan	0	1	2	0
	Civil Penalty	0	1	8	1

R = .97; p < .001

Chart 1a: Tiers By Proposed Actions (Cohort 2)

	Tiers	1	2	3	4
Proposed	**None**	312	0	0	0
Actions	**Tech Assist**	14	43	5	0
	Safety Plan	0	1	2	1
	Civil Penalty	0	4	15	4

R = .80; p < .001

Chart 2: Risk Assessment Matrix (RAM) By Regulatory Compliance (RC) Levels & Licensing Decision Tiers (Cohort 1)

	Tiers	Actions	Immediate	Short Term	Long Term	RC
RAM	.50*	.48*	.63*	.69*	.37*	.93*

* P < .01

Chart 2a: Risk Assessment Matrix (RAM) By Regulatory Compliance (RC) Levels & Licensing Decision Tiers (Cohort 2)

	Tiers	Actions	Immediate	Short Term	Long Term	RC
RAM	.52*	.50*	.62*	.66*	.41*	.88*

* P < .01

Chart 3: Regulatory Compliance Levels By Program Quality Scores (ERS Average Scores)(Cohort 1)

Licensing Bucket	Legend	Compliance	Programs	ERS Aver Score
0	Full	0 violations	33	3.84*
1	Substantial	1-3 violations	32	4.26*
2	Middle	4-9 violations	50	4.18*
3	Low	10+ violations	31	3.92*

* P < .03

Chart 3a: Regulatory Compliance Levels By Program Quality Scores (ERS Average Scores)(Cohort 2)

Licensing Bucket	Legend	Compliance	Programs	ERS Aver Score
0	Full	0 violations	82	4.07*
1	Substantial	1-2 violations	69	4.28*
2	Middle	3-10 violations	163	4.17*
3	Low	11+ violations	71	3.93*

* P < .01

There are some additional significant relationships to report which occurred in the second cohort but were not observed in the first cohort but that was because the total number of sites were fewer in the first cohort. The second cohort had over twice as many sites where data were collected. Here are some of the significant relationships observed between the Quality Rating and Improvement System (QRIS) and regulatory compliance (RC) and the RAM licensing decision making.

- **QRIS x RAM: X^2 = 35.243; p < .009**
- **QRIS x RC: X^2 = 27.761; p < .001**

Significant relationships between Environmental Rating Scales (ERS) and Licensing Decision Tiers (Tiers).

- **ERS x Tiers: F = 5.085; p < .002, where Tier1 = 4.16; Tier2 = 4.10; Tier3 = 3.68; Tier4 = 3.58**
- **ERS x QRIS: F = 26.534; p < .0001, where QRIS1= 3.89; QRIS2= 3.32; QRIS3 = 4.14; QRIS4 = 4.62**

There were interesting demographic and descriptive data with Cohort 2.

- **Regulatory compliance ranged from 0 to 55 violations.**
- **QRIS Levels: 1 = 1%; 2 = 7%; 3 = 78%; 4 = 10%**
- **Licensing Tiers: 1 = 81%; 2 = 12%; 3 = 6%; 4 = 1%**

In both cohorts, there were no significant relationships between regulatory compliance and the CLASS tool as there was with the ERS tool. There was not as much variance in the CLASS tool when compared to the ERS tool. Statistically this was demonstrated when basic distributions were compared and the CLASS's skewness and kurtosis were significantly different than the ERS distribution statistics. These results are consistent with previous studies and warrants additional exploration.

On the basis of the results of this study involving these two independent cohorts, the Washington State DCYF's Risk Assessment Licensing Decision Making Tiers System has been validated at both the measures and output levels. In a previous analysis, the standards that make up the DCYF's Risk Assessment Licensing Decision Making Tiers System have also been validated (see Stevens, 2019 analysis). This state study joins other studies completed which also validated their respective systems core rules & key indicator systems in Georgia and Saskatchewan (see Fiene, 2014 and Fiene, 2020).

The following tables and graphs contain the detail of the above summary analyses and the risk assessment licensing decision making tier system for Cohort 2.

Table 1: Regulatory Compliance: Number of Violations

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	85	21.1	21.1	21.1
	1	43	10.7	10.7	31.8
	2	29	7.2	7.2	39.0
	3	36	8.9	8.9	47.9
	4	27	6.7	6.7	54.6
	5	22	5.5	5.5	60.0
	6	21	5.2	5.2	65.3
	7	23	5.7	5.7	71.0
	8	17	4.2	4.2	75.2
	9	14	3.5	3.5	78.7
	10	11	2.7	2.7	81.4
	11	13	3.2	3.2	84.6
	12	7	1.7	1.7	86.4
	13	8	2.0	2.0	88.3
	14	9	2.2	2.2	90.6
	15	6	1.5	1.5	92.1
	16	4	1.0	1.0	93.1
	17	4	1.0	1.0	94.0
	18	4	1.0	1.0	95.0
	19	3	.7	.7	95.8
	20	1	.2	.2	96.0
	21	1	.2	.2	96.3
	22	1	.2	.2	96.5
	23	2	.5	.5	97.0
	24	1	.2	.2	97.3
	25	3	.7	.7	98.0
	27	2	.5	.5	98.5
	30	1	.2	.2	98.8
	32	1	.2	.2	99.0
	33	1	.2	.2	99.3
	40	1	.2	.2	99.5
	45	1	.2	.2	99.8
	55	1	.2	.2	100.0
	Total	403	100.0	100.0	

The above table (Table 1) provides the frequency distribution for regulatory compliance (NC) for the Washington State ECE sites that were in cohort 2. From the distribution it clearly demonstrates how skewed the data are where the majority of sites (practically 50% of the sites) are either in full or substantial regulatory compliance with Washington licensing rules/regulations.

The following Table (Table 2) puts Table 1 results into the key buckets for regulatory compliance analysis: 1 = Low Regulatory Compliance (11 violations or greater); 2 = Med Regulatory Compliance (3-10 violations); 3 = Substantial (Subst) Regulatory Compliance (1-2 violations); and 4 = Full Regulatory Compliance (0 violations).

Table 2: Regulatory Compliance Buckets

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Low	75	18.6	18.6	18.6
	2 Med	171	42.4	42.4	61.0
	3 Subst	72	17.9	17.9	78.9
	4 Full	85	21.1	21.1	100.0
	Total	403	100.0	100.0	

This grouping of regulatory compliance bucketing becomes very important in subsequent analyses because of the nature of these data. As has been stated earlier in this report, regulatory compliance data when compared to program quality data is not a linear relationship. To be sensitive to the non-linear nature of the data, these buckets or groupings of data become very significant.

Table 3 depicts the Tiered Licensing Decision Making. In Washington State's Tiered Licensing decision Making System 1 = Continued licensing; 2 = Technical Assistance; 3 = Safety Plan; 4 = Civil Penalty.

Table 3: Licensing Decision Making Tiers

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	326	80.9	81.3	81.3
	2	48	11.9	12.0	93.3
	3	22	5.5	5.5	98.8
	4	5	1.2	1.2	100.0
	Total	401	99.5	100.0	
Missing	System	2	.5		
Total		403	100.0		

The majority of programs are recommended for continued licensing (80%), while the other 20% will receive more intervention.

The next table (Table 4) depicts the Risk Assessment Matrix Levels (RAM1-9). The last section of this report provides the specific methodology and how RAM1-9 and Tiers are linked together in the Washington State Licensing Risk Assessment and Licensing Decision Making Tiers System.

Table 4: Risk Assessment Matrix (RAM1-9)

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1.00	92	22.8	22.8	22.8
	4.00	62	15.4	15.4	38.2
	5.00	106	26.3	26.3	64.5
	6.00	62	15.4	15.4	79.9
	7.00	3	.7	.7	80.6
	8.00	27	6.7	6.7	87.3
	9.00	51	12.7	12.7	100.0
	Total	403	100.0	100.0	

It is interesting to note that not all cells of the matrix are filled. RAM2 & 3 have no sites in their cells. This is something that will need further exploration but it appears since these are at the lower risk levels that regulatory non-compliance is less likely.

The next three table (Tables 5-7) deal with the relative risk level of regulatory non-compliance based upon a weighting of the specific rule/regulation. Weights of 8, 7 and some 6 are of immediate concern, while weights of 4, 5 and most 6 are of short term concern, and weights of 1, 2, and 3 are of long term concern.

Table 5: Immediate Concern

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	325	80.6	80.6	80.6
	1	63	15.6	15.6	96.3
	2	12	3.0	3.0	99.3
	3	2	.5	.5	99.8
	6	1	.2	.2	100.0
	Total	403	100.0	100.0	

In 20% of the regulatory non-compliance did the rule/regulation rise to being of immediate concern. Table 6 depicts the non-compliance for the short term rules/regulations. These are rules that are not the highest risk rules but they are not the least weighted rules either. They fall somewhere in between. There is a higher level of regulatory non-compliance with these rules.

Table 6: Short Term Concern

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	94	23.3	23.3	23.3
	1	52	12.9	12.9	36.2
	2	37	9.2	9.2	45.4
	3	35	8.7	8.7	54.1
	4	22	5.5	5.5	59.6
	5	27	6.7	6.7	66.3
	6	27	6.7	6.7	73.0
	7	23	5.7	5.7	78.7
	8	12	3.0	3.0	81.6
	9	15	3.7	3.7	85.4
	10	14	3.5	3.5	88.8
	11	7	1.7	1.7	90.6
	12	5	1.2	1.2	91.8
	13	7	1.7	1.7	93.5
	14	4	1.0	1.0	94.5
	15	4	1.0	1.0	95.5
	16	2	.5	.5	96.0
	17	1	.2	.2	96.3
	19	3	.7	.7	97.0
	20	2	.5	.5	97.5
	21	1	.2	.2	97.8
	22	2	.5	.5	98.3
	24	1	.2	.2	98.5
	25	1	.2	.2	98.8
	26	1	.2	.2	99.0
	27	1	.2	.2	99.3
	35	1	.2	.2	99.5
	37	1	.2	.2	99.8
	47	1	.2	.2	100.0
	Total	403	100.0	100.0	

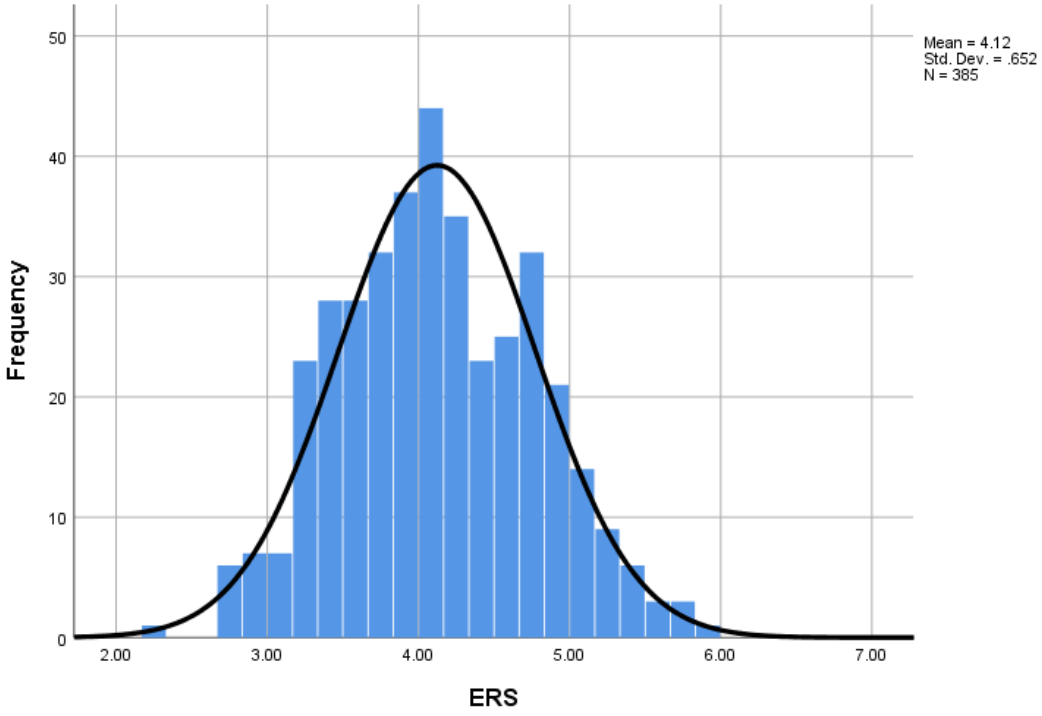
There is a good deal of a range in regulatory non-compliance with these rules as depicted in Table 6. Table 7 which contains the regulatory non-compliance with long term concern rules and regulations which are the lowest weighted/risk rules. The distribution is between the immediate concern and the short term concern rules when it comes to regulatory non-compliance.

Table 7: Long Term Concern

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	224	55.6	55.6	55.6
	1	95	23.6	23.6	79.2
	2	36	8.9	8.9	88.1
	3	21	5.2	5.2	93.3
	4	13	3.2	3.2	96.5
	5	9	2.2	2.2	98.8
	6	1	.2	.2	99.0
	7	1	.2	.2	99.3
	9	1	.2	.2	99.5
	11	1	.2	.2	99.8
	20	1	.2	.2	100.0
	Total	403	100.0	100.0	

The following graphs (Graphs 1-3) depict the distributions of ERS and CLASS scores.

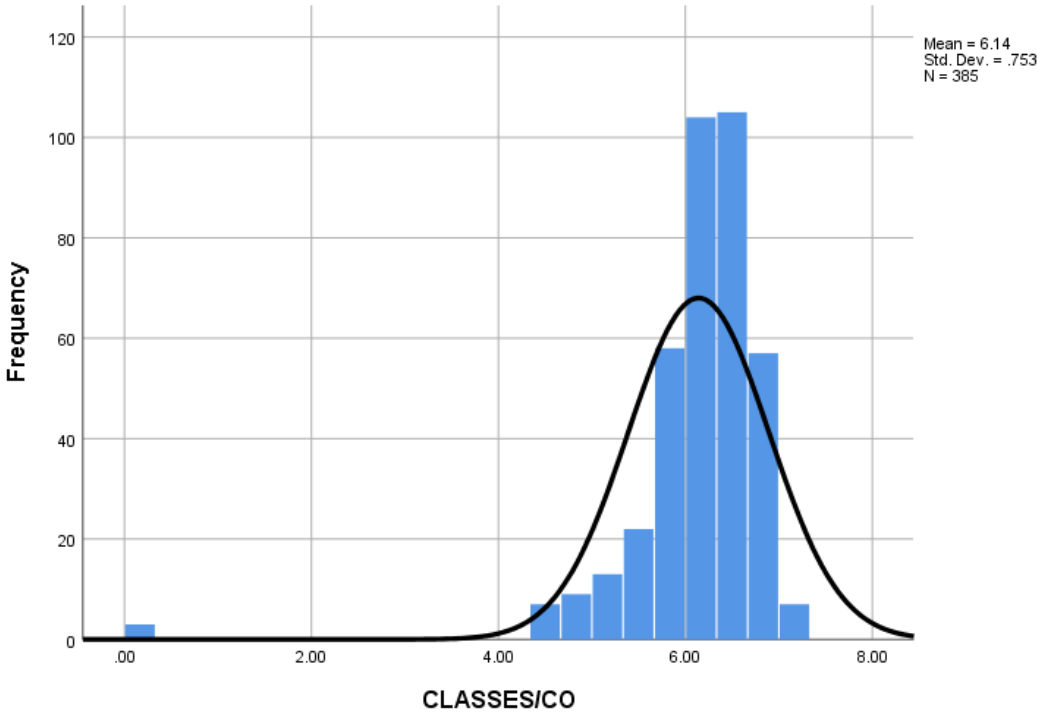
Graph 1 – ERS Scores



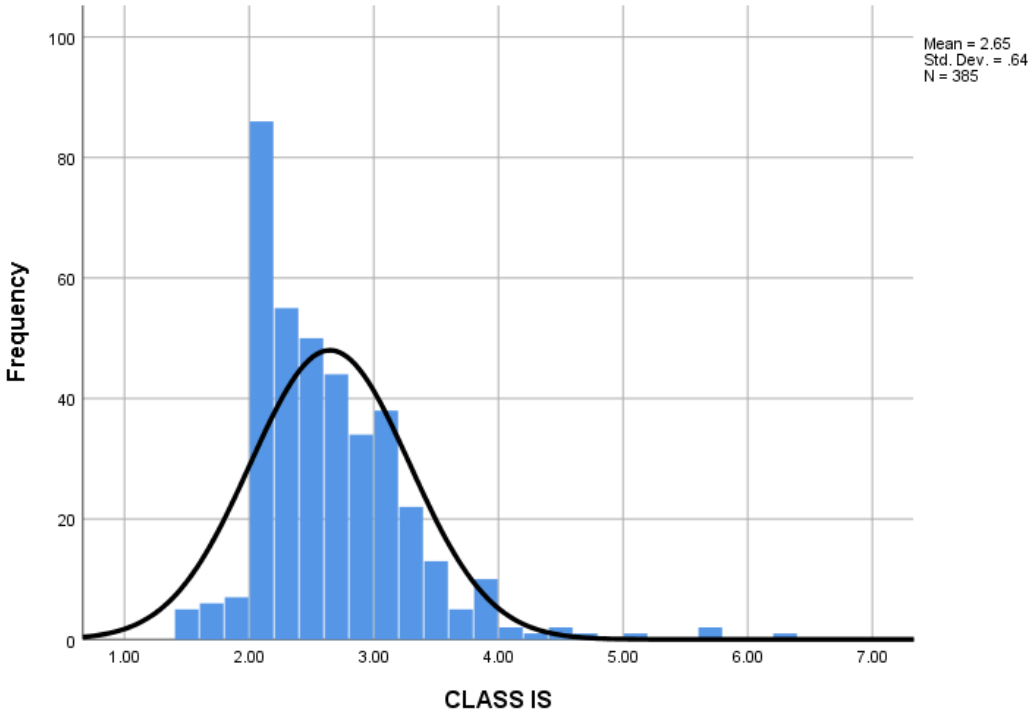
Graph 2 depicts the CLASS/CO scores. Note the difference in the distribution in these scores as versus

the ERS scores in Graph 1. Also note that the N has dropped to 385 sites. This is because not all 403 sites had ERS or CLASS tools administered.

Graph 2: CLASS/CO Scores



Graph 3: CLASS/IS Scores



Again please note the distribution of the CLASS/IS scores and compare it to the CLASS/CO and ERS data score distributions (Compare Graphs 2 & 3 with Graph 1).

Table 8 provides the frequency counts and distribution of the QRIS Levels from 1 to 4 where 4 is the highest level.

Table 8: QRIS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Lowest	2	.5	.5	.5
	2	29	7.2	7.5	8.1
	3	315	78.2	81.8	89.9
	4 Highest	39	9.7	10.1	100.0
	Total	385	95.5	100.0	
Missing	System	18	4.5		
Total		403	100.0		

Table 9 provides the descriptive statistics for all the variables described above so the reader can see the characteristics of the respective data distributions and how they vary.

Table 9: Descriptive Statistics for all Variables

Variables	N	Range	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
NC	403	55	5.93	7.061	2.474	.122	9.739	.243
Immediate	403	6	.25	.592	3.856	.122	24.745	.243
Short	403	47	4.77	5.854	2.640	.122	11.131	.243
Long	403	20	.94	1.720	4.823	.122	40.946	.243
QRIS	385	3	3.02	.445	-.284	.124	3.779	.248
ERS	385	3.64	4.1225	.65207	.120	.124	-.386	.248
CLASSES/CO	385	7.00	6.1411	.75260	-4.514	.124	33.019	.248
CLASS IS	385	4.97	2.6481	.63985	1.658	.124	5.546	.248
RAM1-9	403	8.00	4.8089	2.56860	-.051	.122	-.811	.243
Tiers	401	3	1.27	.617	2.449	.122	5.592	.243
TRC-RCL	403	3.00	2.4144	1.01946	.304	.122	-1.033	.243
Valid N (listwise)	383							

This section describes the Washington State Risk Assessment and Licensing Decision Making Tiered System which was validated in this report.

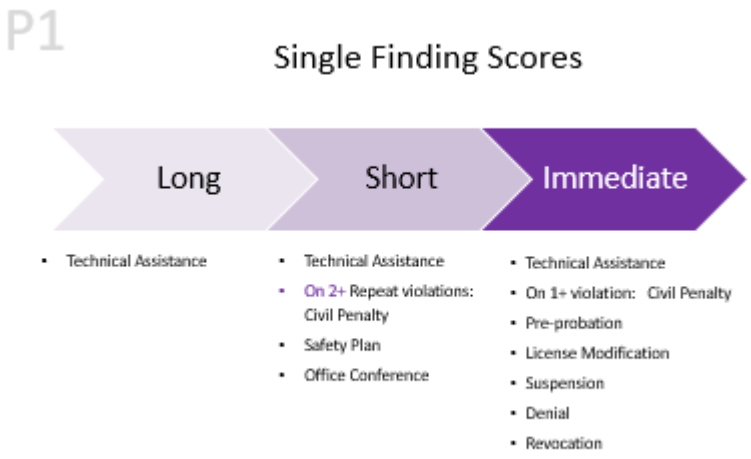
The Washington State System combines the use of risk assessment and licensing decision making matrices. In the past, risk assessment matrices have been used to determine the frequency of monitoring and licensing visits and scope of reviews based upon individual rule severity/risk factors. These data have not been aggregated to determine what type of licensing decisions should be made based upon prevalence, probability or regulatory compliance history data.

Washington State's HB 1661 redesigned the FLCA process as a way to appeal and forgive non-immediate health and safety risks rather than simply being a report of compliance findings. As a result, weights were used to assign risk categories to regulations in accordance to the mandate definition of immediate health and safety regulations:

- **Weights 8, 7 and some 6 = immediate concern**
- **Weights 4, 5 and most 6 = short term concern**
- **Weights 1, 2, and 3 = long term concern**

Single violations of regulations can be considered independently or based on how many time it has been violated over a four-year period when considering licensing actions. For example, a violation within the short term concern category could be subject to a civil penalty when violated the second (or potentially the 3rd) time in a four-year period. Whereas, a violation in the immediate concern category could be subject to a civil penalty or more severe action upon the first violation. (See Graphic for Step 1).

Step 1:



A more difficult task is assigning initial thresholds for the overall finding score. It is this second step (Step 2) where we need to consider probability and severity side by side as depicted in Chart 1 below which is generally considered the standard Risk Assessment Matrix in the licensing research literature:

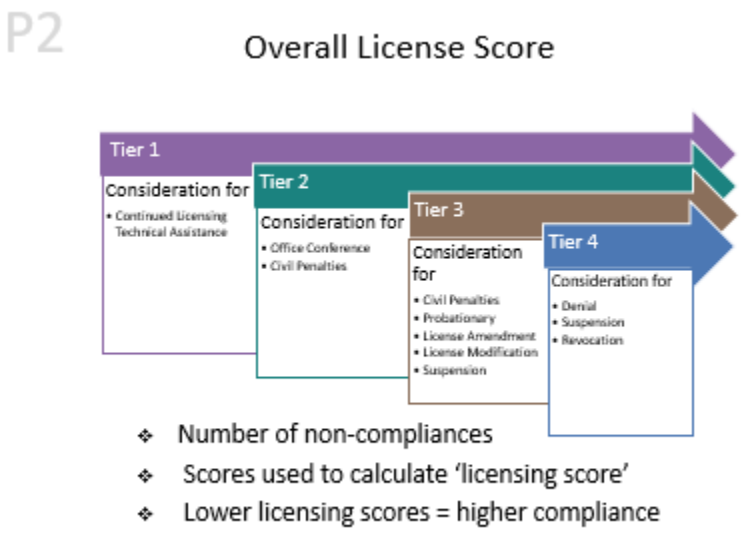
Step 2:

Chart 1 – Risk Assessment Matrix

		Probability/	Prevalence		
	Levels	High	Medium	Low	Weights
Risk/	High	*9*	*8*	*7*	*7-8*
Severity	Medium	*6*	*5*	*4*	*4-6*
	Low	*3*	*2*	*1*	*1-3*
	# of Rules	*8 or more*	*3-7*	*2 or fewer*	

The next step (Step 3) is to build in licensing decisions using a graduated Tiered Level system as depicted in the following figure. In many jurisdictions, a graduated Tiered Level system is used to make determinations related to monitoring visits (frequency and scope) and not necessarily for licensing decisions.

Step 3:



Step 4 involves combining steps 1 and 2 into a revised risk assessment matrix as depicted in the following chart:

Step 4:

Risk Assessment (RA) Matrix Revised

	Levels	High	Medium	Low
Risk/Severity	Immediate	9	8	7
	Short-term	6	5	4
	Long-term	3	2	1
			Probability	
	Regulatory Compliance (RC): # of Rules out of compliance and In compliance	8+ rules out of compliance. 92 or less regulatory compliance.	3-7 rules out of compliance. 93 – 97 regulatory compliance.	2 or fewer rules out of compliance. 98 – 99 regulatory compliance.

The last step (Step 5) is to take steps 3 and 4 and combine them together into the following charts which will provide guidance for making licensing decisions about individual programs based upon regulatory compliance prevalence, probability, and history as well as rule risk/severity data.

Step 5:

Licensing Decision Making Matrix*

Tier 1 = (1 – 2) RA Matrix Score

Tier 2 = (3) RA Matrix Score
Tier 3 = (4 – 5) RA Matrix Score

Tier 4 = (6 – 9) RA Matrix Score

***Regulatory Compliance (RC)(Prevalence/Probability/History + Risk/Severity Level**

Tier 1 = ((RC = 93 – 97) + (Low Risk)); ((98 – 99) + (Low Risk)) = Tier 1

Tier 2 = (RC = 92 or less) + (Low Risk) = Tier 2

Tier 3 = ((RC = 93 – 97) + (Medium Risk)); ((98 – 99) + (Medium Risk)) = Tier 3

Tier 4 = (RC = (92 or less) + (Medium Risk)) = Tier 4; ((93 -97) +(High Risk)) = Tier 4; ((98 – 99) + (High Risk)); ((92 or less) + (High Risk)) = Tier 4+

Key Indicator Matrix (KIM) and Risk Assessment Matrix (RAM)

Key Indicator Matrix

KIM	High Compliance Group	Low Compliance Group
In Compliance	***1***	**2**
Out of Compliance	**3**	***4***

1 + 4 = Key Indicators

2 = False Positives

3 = False Negatives

Risk Assessment Matrix

RAM	High Compliance	Medium Compliance	Low Compliance
Low Risk	***1***	***2***	**3**
Medium Risk	**4**	**5**	**6**
High Risk	**7**	**8**	**9**

1 + 2 = Positive Compliance

3, 4, 5, 6 = Questionable Compliance

7, 8, 9 = Negative Compliance

Fiene, 12/19

The Principles of Regulatory Compliance Measurement

Richard Fiene, Ph.D.

Research Institute for Key Indicators (RIKIIIc)

The Pennsylvania State University

National Association for Regulatory Administration (NARA)

June 2019

The principles of regulatory compliance measurement will be described in this short technical research note covering comprehensive licensing inspections, abbreviated licensing inspections through weighted risk assessment, and how the resultant scoring protocols can be used to make licensing decisions.

Usually when one thinks about regulatory compliance the number of violations are generally the prominent number that most people associate with measuring this concept. So zero (0) violations on a comprehensive licensing inspection is a very good result or number. But what is a not so good number when thinking about regulatory compliance. Based upon the past 40 years of licensing research in which I have established and maintained an international data base related to regulatory compliance, there are trends in data which will help to inform us about what potential thresholds could be in thinking about the number of violations. There is a brief footnote to add to this discussion and that is the impact of the Theory of Regulatory Compliance (Fiene, 1985, 2016, 2019) in which substantial (1-2 violations of low risk rules) and not full compliance (0 violations) is more characteristic of high quality programs.

After taking the Theory of Regulatory Compliance into account, the following ranges based upon the international data base provides us with the following: a provisional level of regulatory non-compliance is between 3 - 7 violations while a low level of regulatory non-compliance is 8+ violations. This results are based upon annual comprehensive licensing inspections in which all rules are measured for compliance. The scoring and license decision making is rather straightforward where if a program has 0 - 2 violations than they would receive a full license; 3 - 7 violations would result in a provisional license with a good deal of technical assistance; and 8+ violations would result in negative sanctions being applied. This scoring protocol takes prevalence data into account but not the relative weight or risk assessment of regulatory non-compliance. That is where differential monitoring can play a role in constructing a licensing risk assessment matrix which is used by a number of jurisdictions in the US and Canada.

Weighted Risk Assessment Matrices have been used to make determinations about individual rules and how often to monitor a program but have not been used in conjunction with License Decision Making as outlined in the above paragraphs. Depicted below is a standard 3 x 3 Risk Assessment Matrix format that is used by the majority of jurisdictions in the US and Canada. In

the more general research literature on risk assessment, the cells may vary from this 3 x 3 format and might use a 4 x 4 or 5 x 5 format, but the result is the same.

Standard Risk Assessment Matrix: Risk Assessment with Probability along the vertical axis and Risk along the horizontal axis

A	**B**	**C**
D	**E**	**F**
G	**H**	**I**

In the above 3 x 3 Risk Assessment Matrix, (A) indicates a very high risk rule with a high likelihood that it will occur, while (I) indicates a very low or no risk rule with a low likelihood that it will occur. (B) through (H) indicate various degrees of risk and probability based upon their position within the Matrix.

Let's merge the risk assessment designation with the regulatory non-compliance probability data from the earlier paragraphs in the following manner: A = (High Risk Rule) + (8+ Violations); B = (High Risk Rule) + (3-7 Violations); C = (High Risk Rule) + (1-2 Violations); D = (Medium Risk Rule) + (8+ Violations); E = (Medium Risk Rule) + (3-7 Violations); F = (Medium Risk Rule) + (1-2 Violations); G = (Low Risk Rule) + (8+ Violations); H = (Low Risk Rule) + (3-7 Violations); I = (Low Risk Rule) + (1-2 Violations).

The last step is now to take the results of the above 3 x 3 Risk Assessment Matrix and combine this with license decision making as was outlined in the above paragraphs for comprehensive inspections. Risk scores are the predominant factor but the probability or prevalence scores do factor into the overall equation in the following manner especially at the high probability levels: A, B, C, D = Negative sanctions; E, F, G = Provisional license; H, I = Full license.

Risk Assessment, Regulatory Non-Compliance and License Decision Making Matrix

A = Negative sanction	**B = Negative sanction**	**C = Negative sanction**
D = Negative sanction	**E = Provisional license**	**F = Provisional license**
G = Provisional license	**H = Full license**	**I = Full license**

By utilizing this matrix a jurisdiction can now account for both risk assessment and regulatory non-compliance data at the same time in order to make a more informed licensing decision. A validation study is being conducted in the state of Washington to determine the effectiveness of these above two matrices (Stevens & Fiene, 2019).

References:

Fiene (1985). Measuring the effectiveness of regulations, *The New England Journal of Human Services*, *5/2*, *pages 38-39*.

Fiene (2016). *The theory of regulatory compliance*, Research Institute for Key Indicators, ResearchGate.

Fiene (2019). Treatise on the theory of regulatory compliance, *Journal of Regulatory Science*, *Volume 7*, pages 1-3.

Stevens & Fiene (2019). *Risk assessment and licensing decision making matrices: Taking into consideration rule severity and regulatory compliance prevalence data*, National Association for Regulatory Administration and the Washington Department of Children, Youth, and Families.

Richard Fiene, Ph.D., Psychologist, Research Institute for Key Indicators (RIKIllc); Professor of Psychology (ret), Penn State University; and Senior Research Consultant, National Association for Regulatory Administration (NARA).

Alberta Child Care
Facility-Based Child Care Quality Indicators
February 24, 2022

Introduction

This document will introduce, for the first time, the concept of Quality Indicators. Quality Indicators use the same methodology employed in designing Key Indicators for licensing regulations.

The Alberta Quality Indicators are based on License Holder Program Plans developed in accordance with Part 1, 6(a),(b) of the Early Learning and Child Care Act. Program Plans are comprehensive documents that encompass key aspects of an early care and education program, including but not limited to developmental needs of children, educational philosophy, interaction with the local community, child guidance, staffing, accident and illness prevention, health care, and supervision policy and practices.

The Quality Indicators can be used in conjunction with the licensing Key Indicators and High-Risk Regulations to develop an efficient, comprehensive approach to License Holder oversight that balances regulatory compliance and child care program quality.

Data Collection

Data was collected through the review of License Holder Program Plans. Reviews were conducted at the regional level between the period May 2021 to November 2021 using a standardized assessment instrument. The assessment instrument included 34 elements, each of which was assigned a score based on the quality of the plan. The score rubric was as follows:

Score	Descriptor	Description
0	Does not meet requirements	No best practice embedded; no confidence in License Holder's (LH) response
1	Does not meet requirements	Very little best practice embedded; low confidence in LH's response
2	Does not meet requirements	Some best practice embedded; some confidence in LH's response
3	Meets minimum requirements	Little to no best practice embedded; moderate confidence in LH's response
4	Meets requirements	Best practice embedded; confidence in LH's response
5	Meets and exceeds requirements	Significant best practice embedded; high confidence in the LH's response

Licensing Officers conducted an initial review of Program Plans, after which technical assistance was provided to License Holders to improve the quality of the Program Plans. Additional Program Plan reviews were conducted after technical assistance was provided by the Licensing Officers. Program Plan review continued until the Licensing Officers were satisfied that the plans were of the highest possible quality, at which point a "final" review was conducted.

The results of each review were then tabulated to produce a dataset, whereupon a "Total Quality Score" (TQS) for each review was obtained by summing the result of each element.

A fundamental concept of differential monitoring is that health and safety is the foundation upon which quality is built. Health and safety are a function of licensing rules, whereas quality is a function of best practices. Elements of quality measurement should not mirror regulatory requirements[1]. Following discussion of the preliminary Quality Indicator findings with Alberta Child Care staff in December 2021 and February 2022, it was determined that multiple elements of the Program Plans are very similar to regulatory requirements in the Early Learning and Child Care Act. As a result of that discussion, the Quality Indicator dataset was truncated to include only Program Plan elements that did not have a corresponding regulatory requirement. The final dataset included the following 12 elements:

Element Number	Description
1	The early learning and child care philosophy the facility-based child care program is based on.
2	How the philosophy will be applied to encourage care and play experiences that support children's development and early learning in the program.
3a	How the child care program plans to meet, promote and nurture the mental needs of children.
3b	How the child care program plans to meet, promote and nurture the emotional needs of children.
3c	How the child care program plans to meet, promote and nurture the spiritual needs of children.
3d	How the child care program plans to meet, promote and nurture the physical needs of children.
4	How the program will be inclusive and accommodate the needs of all children including those with exceptional needs.
5	How the program will incorporate and support the child's familial, Indigenous or other cultural, social, linguistic and spiritual heritage to ensure it is central to the child's safety, well-being and development.
6	How the program will engage with and access community organizations, resources, and members to promote positive connections.
7	The nature and scope of parental involvement in the program
8	Describes the process for ongoing evaluation and improvement of the child care program.
11	Staff orientation.

A new TQS was then obtained using only the above elements.

[1] This long-established concept was first presented in the late 1970s (see Young Children, Vol. 34 No. 6 Sept. 1979, pp. 22-27, Gwen G Morgan) and continuously reinforced since then, e.g. Rick Fiene's December 2012 update of Morgan's research.

Methodology

The methodology sought to measure the strength of the relationship between each individual quality-based Program Plan element and the TQS. The Spearman's Rho test of association was used. The individual element scores are ordinal variables while TQS is a ratio variable, and the variables have a monotonic relationship. For purposes of this analysis, strengths of association were as follows:

Correlation Coefficient	Strength
.00 to .19	Very Weak
.20 to .39	Weak
.40 to .59	Moderate
.60 to .79	Strong
.80 to 1.0	Very Strong

Tests of association were conducted on the initial review and the final review. This method was chosen because the nature of indicators is such that the relationship between individual elements and TQS should be the same regardless of any intervention, i.e., the provision of technical assistance. What we would therefore seek to identify are the elements that have strong and statistically significant associations in both the initial and final reviews. The correlation coefficients and strengths of each element are shown below:

Element	Correlation - Initial	Strength – Initial	Correlation - Final	Strength – Final
1	0.57	Moderate	0.64	Strong
2	0.61	Strong	0.63	Strong
3a	0.78	Strong	0.72	Strong
3b	0.77	Strong	0.70	Strong
3c	0.79	Strong	0.73	Strong
3d	0.82	Very Strong	0.73	Strong
4	0.79	Strong	0.70	Strong
5	0.65	Strong	0.55	Moderate
6	0.78	Strong	0.70	Strong
7	0.70	Strong	0.72	Strong
8	0.58	Moderate	0.63	Strong
11	0.70	Strong	0.67	Strong

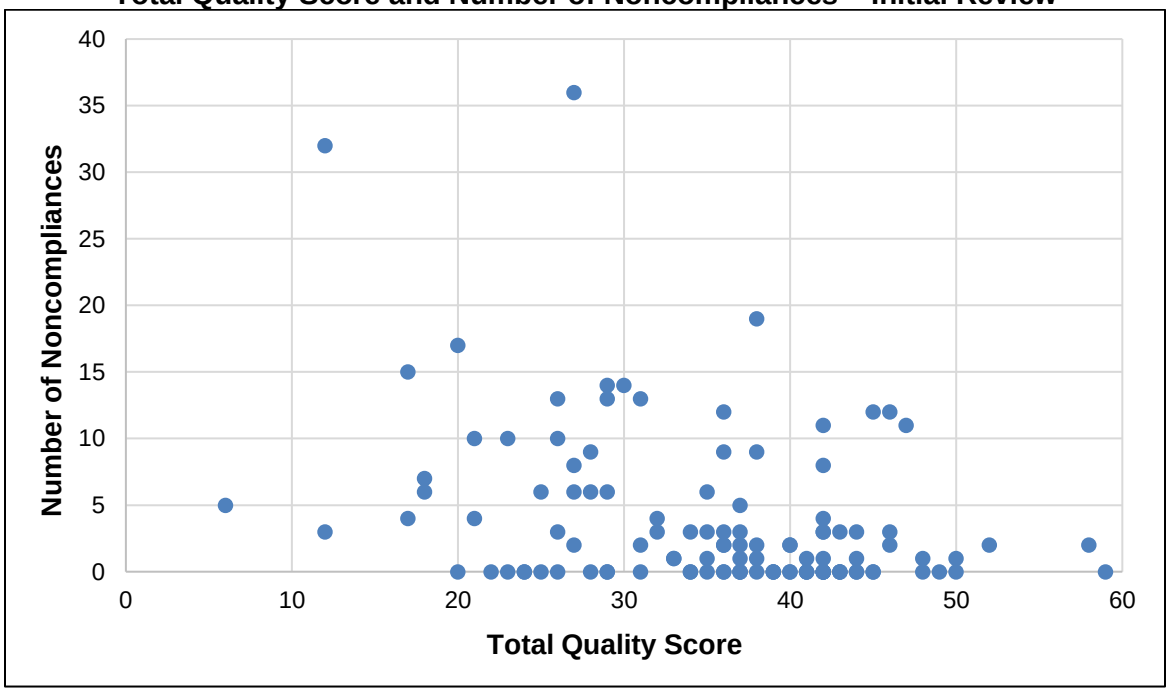
Quality and Noncompliance

Another standard test for the validity of a quality measurement system / Quality Indicators involves measuring the relationship between regulatory compliance and program quality. Historical analyses have consistently found that programs with substantial regulatory compliance have higher quality scores than programs with full regulatory compliance. This is typically called the "regulatory compliance law of diminishing returns."

The Pearson's r test of association was used to determine the relationship between the TQS from both the initial and final reviews and the number of noncompliances identified by each program during the period February 2019 to January 2021.

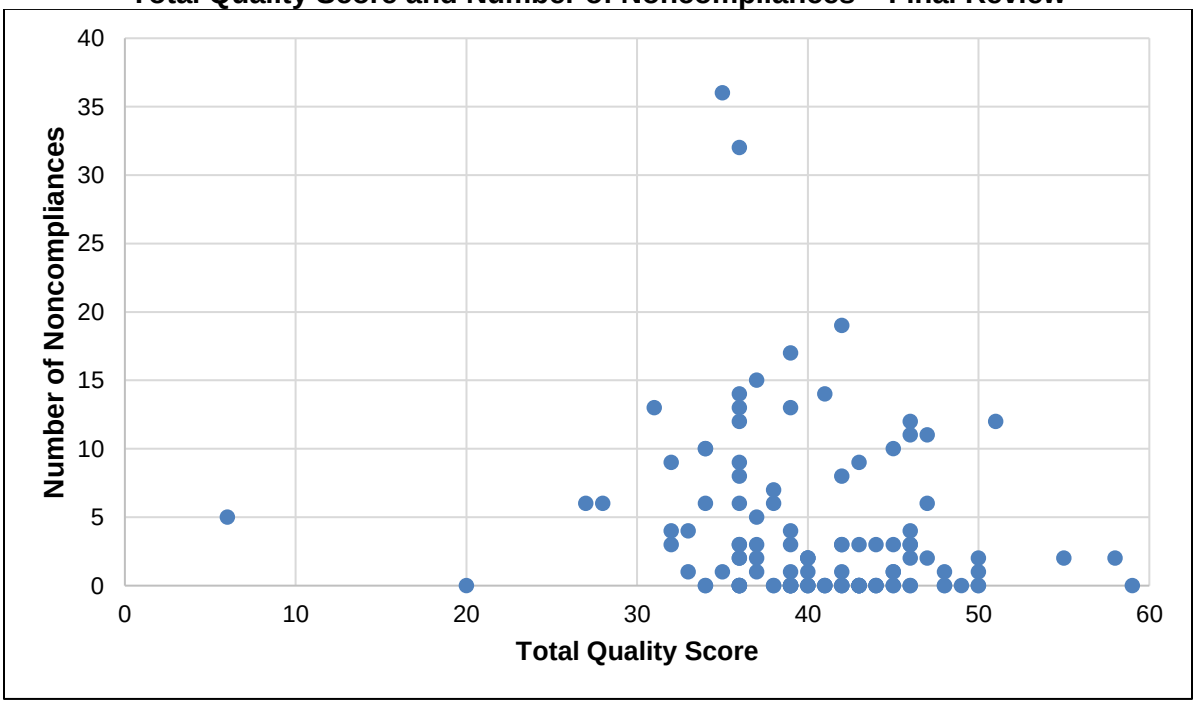
The tests found extremely weak correlations between TQS and the number of noncompliances for either review. The graphs below illustrate the relationship between the TQS and noncompliance.

Total Quality Score and Number of Noncompliances – Initial Review



r = -0.35

Total Quality Score and Number of Noncompliances – Final Review

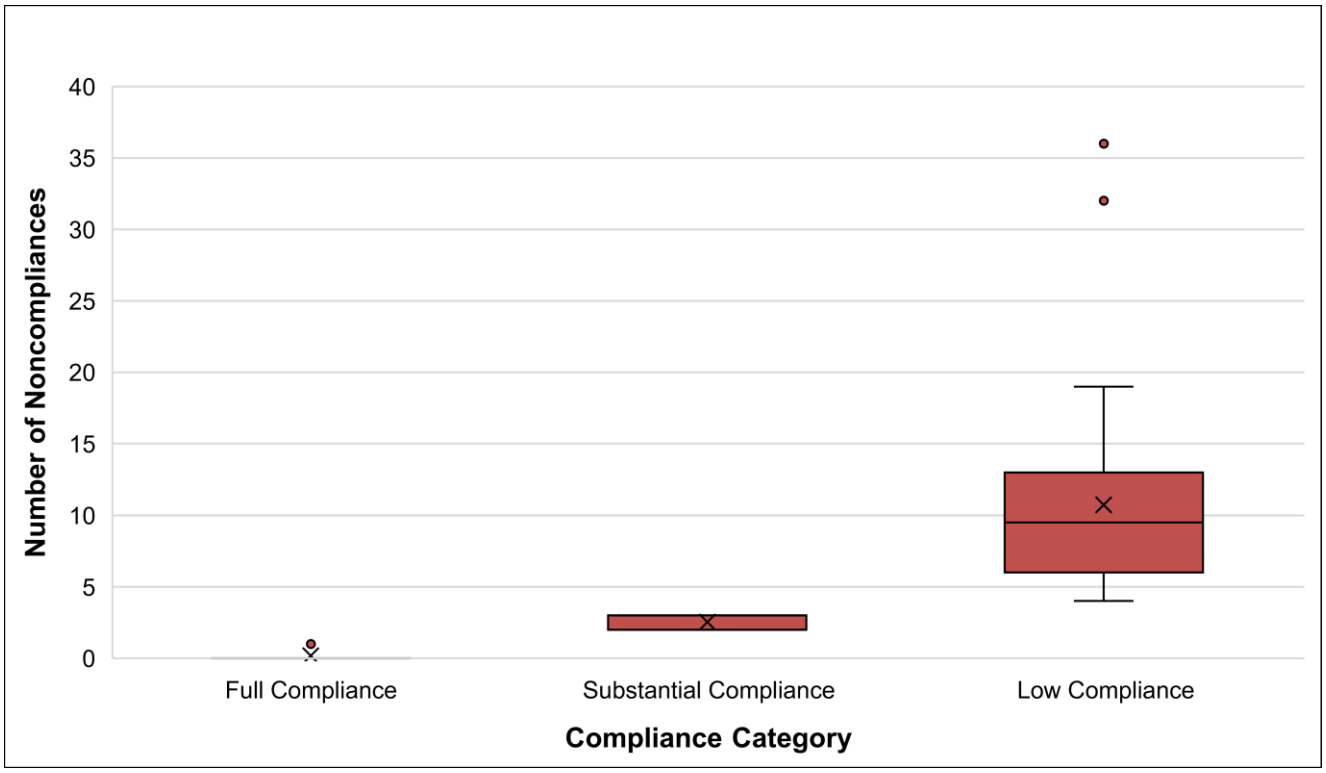


r = -0.18

To further test the relationship between TQS and noncompliance, the license holders were grouped into three categories based on the number of noncompliances:

Compliance Category	Noncompliances	License Holders in Category
Full Compliance	0	44
Substantial Compliance	1-3	33
Low Compliance	4 or more	36

The chart below shows the range of scores in each compliance category, the median or “middle” of the range, and outlier license holders.



A One-Way Analysis of Variance (ANOVA) test was used to compare the three categories to the TQS to determine whether there was a statistically-significant difference between the average TQS of each group that is greater than what one would expect to see by chance. Results:

Initial Review

Compliance Category	Average TQS
Full Compliance	37.4
Substantial Compliance	38.5
Low Compliance	29.1

F = 11.96; p = 0.001

This test shows that there is a statistically-significant difference in the average TQS between each category such that the difference is greater than what one would expect to see by chance. Note that the average TQS for the Substantial Compliance category is higher to the average TQS for the Full Compliance category, which is consistent with the existing literature on quality and regulatory compliance.

Final Review

Compliance Category	Average TQS
Full Compliance	41.4
Substantial Compliance	42.0
Low Compliance	37.3

F = 5.7; p = 0.005

This test shows that there is a moderately-statistically significant difference in the average TQS between each category; such that the difference is greater than what one would expect to see by chance. In this case, the average TQS for the Substantial Compliance category remains higher) than the average TQS for the Full Compliance category.

Results

The results of the above analyses verify the following:

- Each of the 12 elements relating to quality have a "strong" or "very strong" statistically-significant relationship to the overall quality of a child care program, telling us that each quality element is a Quality Indicator.

- The relationship between regulatory compliance and the overall quality of child care programs is consistent with previous research, validating that the 12 quality elements are each Quality Indicators.

Incorporating Quality Indicators into the Differential Monitoring System

Licensing oversight agencies have used a variety of quality evaluation and improvement tools in conjunction with regulatory compliance data for wholistic oversight for many years. However, this is the first time that Quality Indicators have been considered for an abbreviated evaluation of program quality.

One of Alberta's key objectives in their overall Differential Monitoring System is to maximize efficiency of regulatory oversight and quality management during the inspection process. For this reason, NARA recommends that only those Quality Indicators with correlation coefficients of 0.70 or higher for both the initial and final Program Plan reviews be reviewed during inspections. This will reduce the "actual" number of Quality Indicators from 12 to 7:

Element	Description	Correlation - Initial	Strength – Initial	Correlation - Final	Strength – Final
3a	Nurture Mental Needs	0.78	Strong	0.72	Strong
3b	Nurture Emotional Needs	0.77	Strong	0.70	Strong
3c	Nurture Physical Needs	0.79	Strong	0.73	Strong
3d	Nurture Spiritual Needs	0.82	Very Strong	0.73	Strong
4	Inclusivity / Accommodate all Needs	0.79	Strong	0.70	Strong
6	Engage and Access Community	0.78	Strong	0.70	Strong
7	Scope of Parental Involvement	0.70	Strong	0.72	Strong

When performing either full or abbreviated inspections, Licensing Officers will use a Program Plan Evaluation Tool to "rate" the Quality Indicators. The basic process steps will be:

1. The Licensing Officer will ask to see the Program Plan.

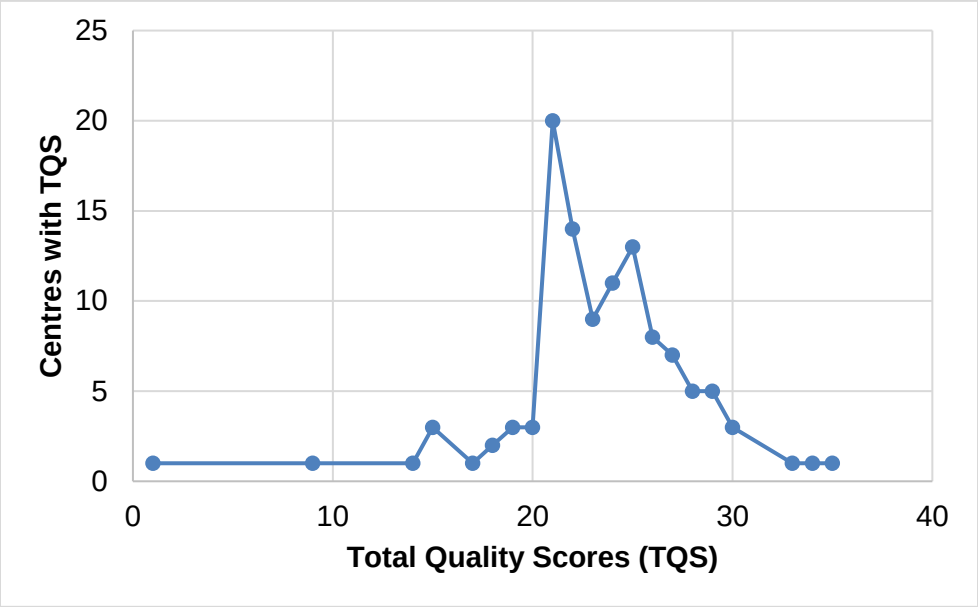
2. The Licensing Officer will use the tool to evaluate 7 Program Plan Quality Indicators (the elements of the program plan) listed above.

3. Each element will be scored by the Licensing Officer using a range of 0 – 5, where 0 is "Unacceptable" and 5 is "Exceptional."

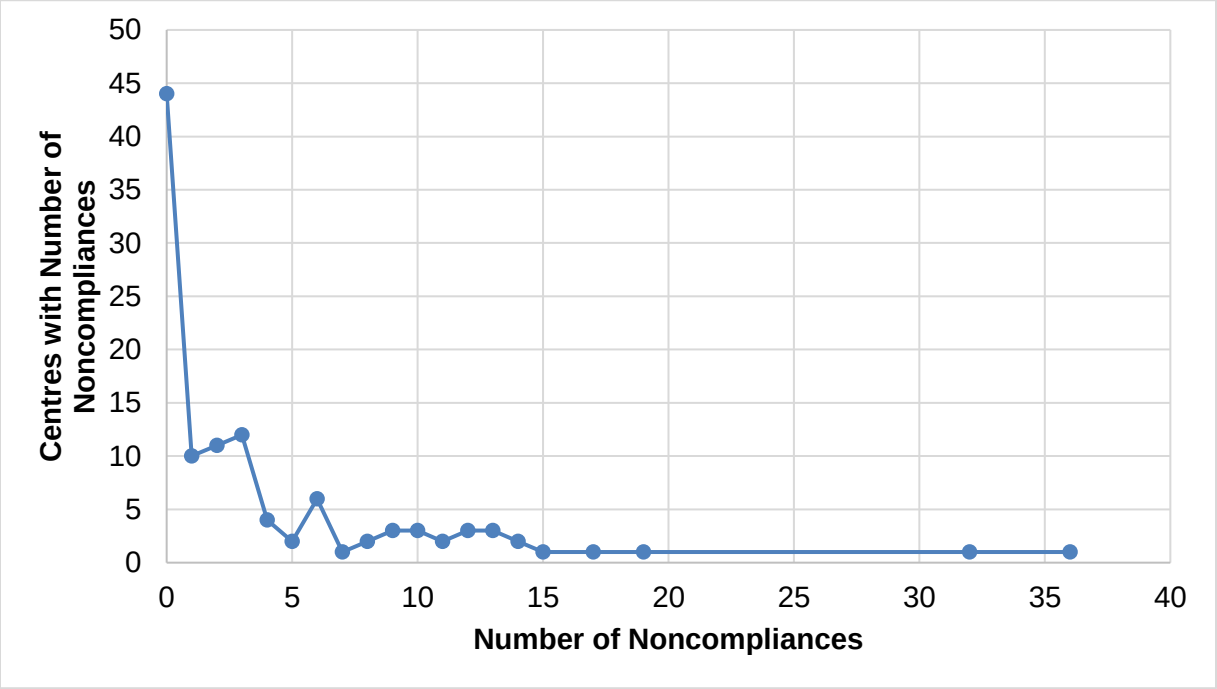
4. The scores for each element will be summed to produce the "Quality Score."
    a. The lowest possible score is 0.
    b. The highest possible score is 35.

Example: Quality Scores from the Quality Indicator Study

The chart below shows the Total Quality Scores for the 113 centres whose Program Plans were reviewed for the Alberta Quality Indicator Study[2]. A small number of programs had very low scores and a small number of programs had very high scores. Most programs had scores between 20 and 25. This is exactly the range of program quality we would expect to see.



The chart below shows number of noncompliances identified during the period February 2019 to January 2021 at the 113 centres whose program plans were reviewed. Most centres had between 0-3 noncompliances. This further illustrates how a center could be highly-compliant with regulatory requirements but not have a high TQS.



[2] The final Program Plan reviews are shown.

How will the Quality Score Be Used?

The Quality Score will be for internal use only. It will be used to:

- Serve as a "quick reference" when there is an immediate need to describe a license holder's performance, e.g., a government official asks "is this a good or bad facility?"
- Provide technical assistance to the License Holder about areas for quality improvement.
- Compare the overall quality of a given centre to other centres.

Conclusion and Next Steps

These findings show that Quality Indicators have been statistically validated and can be used to develop an oversight method that balances regulatory compliance and program quality without compromising the health and safety of children in care. The next step in this process is to test the validity and effectiveness of the method through the differential monitoring pilot.

Appendix – Data Collection and Analysis Information

Table 1: Number of Program Plan Reviews Completed, by Region* (n = 219)

	Number of Reviews						
Region	**1**	**2**	**3**	**4**	**5**	**6**	**Total**
Calgary	23	17	2	1	0	0	43
Central	24	17	3	0	0	0	44
Edmonton	19	13	1	0	0	0	33
North	30	21	2	0	0	0	53
South	17	13	8	4	3	1	46
Total	**113**	**81**	**16**	**5**	**3**	**1**	**219**

Example: Calgary completed 1 review of 23 License Holders' Program Plans, 2 reviews of 23 License Holders' Program Plans, etc.

Table 2: Number of License Holders with at Least 1 Program Plan Review, by Region (n = 113)

Region	License Holders
South	17
Calgary	23
Central	24
Edmonton	19
North	30
Total	**113**

NARA

The Saskatchewan Early Care and Education Quality Indicators Tool and Validation

Abstract

This validation study involved 30 programs, 90 classrooms and 180 observations of infant, toddler, and preschool classrooms utilizing the ECERS/ITERS and the SKECPQI instruments. Six trained observers collected the data over a two-month period. The analyses clearly demonstrated that the new SKECPQI instrument is a valid and reliable measure of program quality. PQI #2 clearly showed it predictive power in this study. The SKECPQI and PQI #2 correlated very highly with the ITERS and ECERS. The SKECPQI appears to correlate more highly with regulatory compliance violations than the ECERS or ITERS. The ceiling/plateauing effect is not as evident with the SKECPQI as it is with ECERS/ITERS. The Regulatory Compliance Scale (RCS) is a better sorter for regulatory compliance than the violation data. There is a good deal of internal consistency within the SKECPQI Tool just as it is with the ERSs. The Regulatory Compliance Theory of Diminishing Returns was validated in comparing RCS with ECERS/ITERS. Both the SKECPQI Scale and the Regulatory Compliance Scale are introduced as new improvements to measuring quality and regulatory compliance.

The Saskatchewan Early Care and Education Quality Indicators Tool and Validation: The Last Piece of the Puzzle in Creating a Differential Monitoring Approach

National Association for Regulatory Administration

May 2023

INTRODUCTION

This report will delineate the development, piloting and validating of the Saskatchewan Early Care and Education Quality Key Indicators (SKECPQI) Tool. The purpose of the tool is to assess the overall program quality in centered based childcare programs in the Province of Saskatchewan, Canada. The evolution of the tool resulted from a multi-year effort by the Ministry of Education in the Province of Saskatchewan to build an effective and efficient differential monitoring system.

This effort in building a new differential monitoring system started in 2019 and was completed in 2023. The first component of this restructuring was the Saskatchewan Licensing Key Indicator System (2019). This was followed by the Saskatchewan Risk Assessment Rules (2019). Once these were in place and operational, a validation study was conducted to measure that the two methodologies were operating as they should (2020). A work group was initiated in 2019 and completed its work in 2020 on an Early Care and Education Quality Key Indicator Tool (SKECPQI). The tool was put on hold for 2021 because of the pandemic and a new Canadian Federal initiative to expand childcare services across the province. The tool initiative began again in 2022. The pilot testing and validation occurred in 2023.

The work and these studies in the Province of Saskatchewan by the Ministry of Education is the first demonstration of a full-blown differential monitoring system involving licensing key indicator rules, risk assessment rules, and quality indicators. Besides the development of each tool, each of these tools have been validated as well. All this work was done as a collaborative effort between the Ministry of Education staff and the National Association for Regulatory Administration (NARA) consultant pool. Presently, Saskatchewan's overall system is the best example of a fully developed differential monitoring system for the early care and education field.

This was a monumental effort involving many individuals at the local, provincial, and national levels and many hours of data collection and analysis. All the reports are available on the NARA Website (https://www.naralicensing.org/key-indicators) and the full data set will be available via Mendeley Data Sources (https://data.mendeley.com/datasets/kzk 6xssx4d/1).

BACKGROUND HISTORY

This study and tool grew out of an interest by Saskatchewan Ministry of Education policy makers to establish a balance between regulatory compliance and program quality in the most effective and efficient manner. The Province of Saskatchewan did not have a QRIS (Quality Rating and Improvement System) in place nor plans on developing one. Generally, when a jurisdiction wants to develop a balance between regulatory compliance and program quality with rules/regulations/standards, QRIS's are generally developed and implemented.

In reviewing the research literature on regulatory science, differential monitoring has been a developing approach used by many other jurisdictions in the human service licensing field, especially in the United States and in several other Canadian Provinces. Based upon this review of the research literature and the work of the National Association for Regulatory Administration (NARA) which has been a long-term promoter of this approach and the resulting methodologies of licensing key indicators, risk assessment rules, and most recently quality indicators, a contract was entered into between the Ministry of Education and NARA.

The tool is the direct result of research into identifying licensing and quality key indicators over a 50-year (1970-2022) research effort in which specific methodologies were developed and the differential monitoring approach was tested and implemented in the 1970's. Since that time, a national database which expanded to an international database of common key indicators from jurisdictions' respective key indicator tools. These key indicators resulted in a very similar tool that Saskatchewan is using. In fact, in 2019 when the Saskatchewan work group was established, they started with that specific tool that had been developed (Fiene, 2019). During the 2019-2020 period, the work group made the tool into a more user-friendly tool for Saskatchewan childcare programs.

The big deal with utilizing the key indicator methodology is its ability to statistically predict as if one administered the full tool in question. Therefore, when one administers the first quality indicator in the Saskatchewan Early Care and Education Quality Indicator tool, it is as if they have administered a licensing based regulatory compliance instrument since the quality of staff is a statistically predictive rule (Fiene, 2002a). The same is true in administering the curriculum quality indicator because it is a statistically predictive standard when looking at overall program quality (Fiene, 2002b). When it comes to QRIS, having communication between staff and parents and parental involvement is a statistically predictive standard for an overall set of QRIS standards (Fiene, 2014). And finally, when administering the ECERS and ITERS or the CIS quality item indicators these are all statistically predictive items for their respective scales as if you had administered the full scales (Fiene, 2002b).

So, as a state/provincial administrator, I would be interested in focusing my efforts on these indicators which reflect compliance with high quality rules/regulations/standards for early care and education. This would be my starting point. I would make sure that my standards reflected quality teachers with the necessary supports such as coaching/mentoring, an early care and education philosophy based upon an emergent curriculum where children are viewed as competent learners, developmentally appropriate curriculum and child assessments, parental and staff communication and participation, and teacher language based/communicative focus when interacting with children in a give and take manner. All this done within a warm and loving style.

An even more efficient and effective way of using the new program quality tool is to pair it with the National Center for Health and Safety in Child Care's *Parental Guide to Choosing Safe and Healthy Child Care (DHHS: Assistant Secretary's Office for Planning and Evaluation, 2019)*. This is a more aggressive and controversial approach, but it is the most efficient way of conducting monitoring visits in the most abbreviated way. However, as efficiency increases, effectiveness may decrease; so, it is a delicate balancing act. This suggested approach builds off a similar suggestion in which only using *Caring for Our Children: Basics (ACF, 2015)* a DHHS Administration for Children and Families publication would be used as the base for regulatory compliance in the United States.

Differential monitoring grew out of a need for jurisdictions to be more effective and efficient in their oversight and inspection efforts of early care and education programs. This started to occur in the late 1960's and 1970's as many more programs were being established. It was becoming clear that the old one size fits all approach to program monitoring was being overwhelmed by the increasing numbers of programs. Also, from an efficiency standpoint it did not make sense to spend the same amount of time with programs that were performing well as those that really needed additional attention. The birth of differential monitoring occurred which at that time it was called inferential inspections (Fiene & Kroh, 2000). Different terminology, same concept.

Since then, differential monitoring has two basic methodologies that have been used successfully over the years: risk assessment and key indicators. The two methodologies have the same results, shortened or abbreviated reviews but they differ in their approaches. Risk assessment as the name implies identifies specific standards that place clients/children at greatest risk or morbidity or mortality if not complied with. Key indicators are specific standards that statistically predict overall regulatory compliance with all rules. Each has their place in the differential monitoring approach depending on the jurisdictions' emphasis. Most recently, to balance the emphasis on regulatory compliance has been the introduction of quality indicators which are specific standards drawn from quality initiatives, such as professional development, program quality tools, and quality rating & improvement systems.

It is and always has been recommended that these methodologies be used together and not separately. This final study undertaken in the Province of Saskatchewan completes the cycle of doing just that in developing a fully functional differential monitoring system with key licensing and quality indicators as well as risk assessment rules.

THE STUDY DESIGN AND METHOD

The design of this study was to provide a validation study of the use of the Saskatchewan Early Care and Education Quality Key Indicators Tool. A convenience sample was selected in which a good variation of overall quality would be present. There were to be three buckets of quality: High, Middle, and Low. These would be defined via ERS scores. Because this was a validation study it was critical to have sufficient variation in the overall quality of programs to test the sensitivity of the new assessment tool.

The below table (Table 1) provided the guidance to the Saskatchewan Ministry of Education policy staff in determining how to collect the program quality data for the research pilot study related to early childhood quality indicators.

Table 1: Selection Process for Study Programs

Quality	Centers	Classrooms	Ages	Levels	ERS	SKECPQI
High	10	30	10	Infant	A	1
			10	Toddler	B	2
			10	Preschool	C	3
Middle	10	30	10	Infant	A	1
			10	Toddler	B	2
			10	Preschool	C	3
Low	10	30	10	Infant	A	1
			10	Toddler	B	2
			10	Preschool	C	3

Notes:

A = ITERS (Infants) (B-1yr)
B = ITERS (Toddlers) (1yr-2yrs)
C = ECERS (Preschoolers) (3+yrs)
1 = SKECPQI/Infant (QI items 1-5, 7, 9-10)
2 = SKECPQI/Toddler or Preschool (QI items 1-5, 7, 9-10) or (QI items 1-6, 8-10)
3 = SKECPQI/Preschool (QI items 1-6, 8-10)
SKECPQI = Saskatchewan Early Childhood Program Quality Indicators tool

A total of 6 trained data collectors were needed, 3 for the ERSs and 3 for the SKECPQI. Each observer collected data from 30 classrooms. A data coordinator was utilized who collected all the data, reviewed the scores from the various tools and sent them to Dr Fiene. The data collectors were not aware of which centers are in which group, such as High, Middle, or Low

See the Appendix for the Draft of the SKECQKI tool that was used during data collection.

As said earlier, this study involves the validation of the Saskatchewan Early Childhood Quality Indicators Tool (SKECPQI) and involved the collection of new data utilizing the new tool and collecting Early Childhood Environmental Rating Scale (ECERS/ITERS) data as well. Independent contract staff were trained in the use of the SECQIT as well as having had training on the ECERS/ITERS and were proficiently reliable on the ECERS/ITERS.

A sample of 30 childcare programs who volunteer to be part of this study was selected with 1/3 identified as high quality, 1/3 identified as medium quality, 1/3 identified as low quality. Each program had both the SKECPQI and the ECERS/ITERS administered to them utilizing two independent observers. The data from the SKECPQI was compared to the ECERS/ITERS to determine the relationship between the two/three scales. The research hypothesis is that there will be a positive relationship between the two/three scales in which those programs that score high on the SKECPQI will score high on the ECERS/ITERS and those that score low on the SKECPQI will score low on the ECERS/ITERS. The ECERS/ITERS will be used as the reference tool for establishing the validity of the SKECPQI.

A training program and all necessary revisions to policies and procedures was conducted as part of this project by a NARA Consultant on both phase 1 and 2. It will be determined later if the SKECPQI will be administered on an ongoing basis by contracted staff or by Ministry staff. Reporting templates were

developed as part of this implementation stage. The implementation stage was evaluated to make certain that all components are in place and working as they should.

Timeline: Phase 1: 6 months; Phase 2: 9 months; Training and Implementation Phase: 12 months, will overlap with phase 1 and 2 and extend beyond both. The total time frame will be 24 months (about 2 years), this will include the final report and final evaluation of the implementation stage

RESULTS

The ECERS and ITERS were used to validate the new Saskatchewan Early Care and Education Quality Indicators Tool. This is standard procedure when conducting a validation study, a recognized empirically based and accepted standard tool is used in correlational analyses to determine if the new tool is measuring the same dimensions as the standardized tool.

The target tool, the Saskatchewan Early Care and Education Quality Indicators, was to be validated against the ECERS and ITERS to determine if there was a quality relationship between the two tools.

The validation analyses involved detailed correlational analyses between the various scales to determine if a relationship existed and how strong that relationship was. But before delving into this relationship and these analyses, an additional analysis was performed given the sophisticated nature of the Saskatchewan monitoring system. Saskatchewan's Ministry of Education's designed differential monitoring system is by far the most analyzed of all jurisdictions to date, so it was suggested to take advantage of this level of detail and build in an additional series of analyses to further test the regulatory compliance theory of diminishing returns in conducting this study. By doing so, Saskatchewan joins the ranks of the Provinces of Alberta and Ontario, the US States of Georgia and Washington, and the US National Head Start program in conducting studies to either confirm or not this theory of regulatory compliance (please see the NARA website on key indicators which contains all the research reports). The following results delineate the data from that portion of the study.

As part of the data collection in addition to collecting data on the ECERS and ITERS as well as the Saskatchewan Early Childhood Program Quality Indicators scale, a summary sheet containing regulatory compliance data was also obtained on each program. These data contained essential demographic information as well as violations from the last inspection along with a rating of the program which was cross referenced to the regulatory compliance data to generate a Regulatory Compliance Scale. This Regulatory Compliance Scale (RCS) had four levels of regulatory compliance: Full, Substantial, Medium, and Low. This RCS is like the regulatory compliance structure used in the previous studies in the above-mentioned jurisdictions in the US and Canada and has been further developed as a more valid means for measuring and analyzing regulatory compliance (Fiene, 2022). In the Fiene RCS, the following rubric was used: ***Full = 0 violations; Substantial = 1-3 violations; Medium = 4-9 violations; and Low = 10+ violations.***

The first set of analyses was to determine if a correlation existed between the RCS and the ECERS and ITERS. This was the case with the following results: RCS x ITERS for the infant classrooms = .54; $p < .002$; RCS x ITERS for the toddler classrooms = .42; $p < .03$; and RCS x ECERS for the preschool classrooms = .75; $p < .0001$.

The second level of analyses (ANOVA) was to determine if the RCS levels of Full, Substantial, Medium, and Low demonstrated any significant differences in the ECERS and ITERS. The results were the

following: Infant classrooms: Low = 3.07; Medium = 4.89; Substantial = 5.06; Full = 4.69; F = 11.43; p < .0001. Toddler classrooms: Low = 3.50; Medium = 4.56; Substantial = 4.62; Full = 5.06; F = 2.27; p < .11. Preschool classrooms: Low = 2.78; Medium = 4.39; Substantial = 4.90; Full = 5.12; F = 16.27; p < .0001. Apart from the toddler classrooms, both the infant and preschool classrooms support the regulatory compliance theory of diminishing returns ceiling and plateauing effect when it comes to measuring program quality as one moves up the regulatory compliance scale.

Table 2: Regulatory Compliance Scale (RCS) and ECERS/ITERS Scores

RCS	Infant Classrooms	Toddler Classrooms	Preschool Classrooms
Low	3.07	3.50	2.78
Medium	4.89	4.56	4.39
Substantial	5.06	4.62	4.90
Full	4.69	5.06	5.12
Significance	F = 11.43; p < .0001	F = 2.27; p < .11 NS	F = 16.27; p < .0001

ECERS, ITERS for Infant classrooms, ITERS for Toddler classrooms (n = 90):

The ECERS score ranged from 1.41 to 6.00. The ITERS for infant classrooms ranged from 2.16 to 5.77; and the ITERS for toddler classrooms ranged from 2.14 to 5.90. The respective means for the ECERS, ITERS-Infant classrooms, and the ITERS-Toddler classrooms were the following: 4.09, 4.39, 4.39. The means and ranges were all consistent.

The correlations of the infant, toddler and preschool classrooms in each of the 30 facilities were the following: Infant and Toddler classrooms = .65; p < .0001; Infant and Preschool classrooms = .74; p < .0001; and Toddler and Preschool classrooms = .52; p < .005. The classrooms demonstrated a great deal of consistency across the various facilities which one would expect.

SKECPQI for Preschool, Infant, and Toddler Classrooms (n = 90):

The SKECPQI score ranged from 13 to 100. The SKECPQI for infant classrooms ranged from 33 to 91; the SKECPQI for toddler classrooms ranged from 13 to 72; and the SKECPQI for preschool classrooms ranged from 25 to 100.

The correlations of the infant, toddler, and preschool classrooms in each of the 30 facilities were the following: Infant and Toddler classrooms = .73; p < .0001; Infant and Preschool classrooms = .85; p < .0001; and Toddler and Preschool classrooms = .74; p < .0001. The classrooms demonstrated a great deal of consistency across the various facilities which one would hope to be the case with this type of tool or scale. Based upon these results, the inter-correlations were extremely high and show a great deal of stability and are a reliable measure of quality indicators.

SKECPQI #2 showed a great deal of promise as a standalone quality indicator. SKECPQI#2 correlated significantly with ITERS (.53; p < .0001), and ECERS (.61; p < .0001) and with the overall SKECPQI scores for infant classrooms (.87; p < .0001), toddler classrooms (.82; p < .0001), and preschool classrooms (.90; p < .0001). This quality indicator dealt with philosophy, curriculum planning and programming. This is not the first time that such an indicator was an excellent predictor. This result has been the case in other program quality studies as well (Fiene, Greenberg, Bergsten, Fegley, Carl, Gibbons, 2002b).

The SKECPQI scale demonstrated a great deal of robustness in the data distribution and a good deal of variation in the data set. These are the characteristics of a new tool that you would hope to find in the scale construction and implementation.

Regulatory Compliance Data for Each of the Programs (n = 30):

The Regulatory Compliance Scale (RCS) distributions were the following: Full = 13%; Substantial = 20%; Medium = 37%; and Low = 27%. Generally regulatory compliance data are more skewed than this distribution but because of the nature of this study, facilities were deliberately selected breaking them up into these categories/levels.

The Regulatory Compliance Scale (RCS) actual regulatory compliance violations played out in the following table, these results for the average number of violations were statistically significant ($F = 3.69$; $p < .03$):

Table 3: Regulatory Compliance Scale by the Number of Violations

RCS	Regulatory Compliance Means	Number of Facilities
Low	4.75	8
Medium	3.90	10
Substantial	1.60	5
Full	0	4

Comparing the ECERS and ITERS with SKECPQI and Regulatory Compliance (RCS) Data:

These are the correlations between RCS and SKECPQI for infants, toddlers, and preschool classrooms. RCS x PQI for the infant classrooms = .54; $p < .004$; RCS x SKECPQI for the toddler classrooms = .46; $p < .002$; and RCS x SKECPQI for the preschool classrooms = .58; $p < .002$. The SKECPQI clearly demonstrates its relationship with regulatory compliance. Also, when the SKECPQI is compared with regulatory compliance violation data, the correlations are higher than those obtained in comparing the ERSs to regulatory compliance violation data. And, in fact, the SKECPQI when compared with the RCS appears not to have a ceiling or plateauing effect. It would appear that the SKECPQI is measuring quality in a different way since this effect does not appear evident in the RCS distributions. This result will need to be confirmed in other studies to make certain this relationship holds up. This is a first for comparing regulatory compliance data with program quality data. In the past, either a ceiling or plateauing effect was always present when looking at the relationship between regulatory compliance and program quality.

Here are the correlations between SKECPQIs and ERSs for infant, toddler, and preschool classrooms: PQI x ITERS for the infant classrooms = .59; $p < .001$; PQI x ITERS for the toddler classrooms = .50; $p < .009$; and PQI x ECERS for the preschool classrooms = .64; $p < .0001$. These inter-correlations most definitely suggest that the SKECPQI is a valid tool measuring program quality on a different dimension (quality indicators) than the ERS but measuring quality, nonetheless.

A regression analysis determined that with RCS as the dependent variable, ECERS and regulatory violations were statistically significant at the $p < .0001$ with an $R = .91$. This accounted for practically 75% of the variance in being able to determine regulatory compliance.

DISCUSSION

Last piece of the puzzle in creating a differential monitoring system, that is how this report is being characterized. The Province of Saskatchewan has undertaken all the other methodologies utilized in a differential monitoring approach (Please see the NARA website for these reports, the link is hot linked on the first page of this report). Licensing key indicators and risk assessment rules have been implemented successfully. What remained were the Quality Indicators. This report completes the full cycle of validating these last indicators.

With the completion of this validation study, the Saskatchewan Early Childhood Program Quality Indicators Scale could be adapted by other jurisdictions and utilized as a screener methodology. The reason for suggesting this approach is that all the quality indicators are taken from the Key Indicator Methodology and therefore have predictive value when it comes to determining overall quality (Fiene, 2019a). Also, the indicators are drawn from several early care and education delivery systems and quality initiatives, such as licensing, QRIS, quality scales, accreditation, and professional development.

The other significant finding from this study was the additional confirmation of the regulatory compliance theory of diminishing returns in which the results from this study are consistent with the findings from other studies conducted in Canada and the United States. This continues to be a major finding when it comes to comparing regulatory compliance with program quality and the resulting ceiling and/or plateauing effect related to quality scores. Again, from a public policy viewpoint, this finding has significant implications in how licensing decisions are or should be made.

A very interesting finding which was not expected was the fact that when the SKECPQI scores were compared with the regulatory compliance violation data the usual ceiling/plateauing effect did not emerge as in previous studies when these types of analyses were performed. This result needs further exploration to determine why this occurred. In future studies utilizing the SKECPQI, it will be necessary to do similar analyses with regulatory compliance data to ascertain if this same result occurs. At this point, it is difficult to determine if it is characteristic within the SKECPQI that is producing this result, such as a better balance between regulatory compliance and program quality. Only with further study will we be better able to determine the cause of this different result.

CONCLUSION

I am sure that this report will be read with a certain amount of skepticism in that it suggests using differential monitoring on a much broader scale; however, this report is like several other validation studies conducted by NARA over the past decade which have now clearly demonstrated the validity of the differential monitoring approach. And because of these validation studies, the differential monitoring approach has been utilized by many jurisdictions and has been cited in the United States Federal Legislation that reauthorized the Child Care and Development Block Grant. In the legislation, it is suggested but not required that states entertain the use of the approach. Based upon the latest childcare licensing data, it appears that many states have attempted to utilize the approach.

This report fits with the other regulatory compliance theory reports from states and provinces that have been completed over the past decade by NARA. As mentioned in the **Results and Discussion Sections**, this study is the most comprehensive of the group since the Province of Saskatchewan developed not

only risk rules and key indicator rules for licensing but also quality indicators that could be used within their differential monitoring system. This is the first demonstration of this comprehensive approach.

This study and report complete what was to be a three-year effort but turned into a five-year effort because of the COVID19 Pandemic. Each component of this overall project is well documented on the NARA Key Indicator website. The three major results of this study: confirmation of the regulatory compliance theory of diminishing returns, the introduction of the regulatory compliance scale and the introduction of the Saskatchewan Early Childhood Program Quality Indicators Tool/Scale are all significant contributions to the licensing research literature, but it is this last contribution that needs further development.

The Saskatchewan Early Childhood Program Quality Indicators Tool/Scale is a new program quality tool that is rather robust in measuring quality using key indicators which are taken from various quality initiative studies conducted over the past several decades. The hope is that it will continue within the early care and education field being validated by other researchers and being used to determine the relative scope of program quality in various early care and education settings. I could see the scale being utilized throughout the United States and Canada. It would be an excellent supplement to either the ERS or CLASS tools. It is a simple, straightforward tool that can be easily trained on and administered. It could provide an interesting supplement for licensing staff when they are doing their licensing reviews. In fact, it is intended to be used in conjunction with licensing key indicators and risk rule tools.

Although this was not reported in the **Results Section**, I think it is vitally important to highlight the significant contributions of the licensing staff and others who helped to develop the groupings and levels of regulatory compliance and quality. It was only because of their level of early childhood expertise and their knowledge of the programs that made the sequencing so effective and impactful as an analytical frame of reference.

One last thought is the introduction of the Regulatory Compliance Scale (RCS) as a more logical and robust rubric when comparing regulatory compliance data with program quality. This thought has been presented elsewhere as a possible improvement within licensing measurement and monitoring systems (Fiene, 2022). The scale has been piloted in the past, but this is the first formal test of it in a specific jurisdiction.

NOTES:

Saskatchewan Ministry of Education Program Quality Tool Work Group: Kim Taylor, Derek Pardy, Cindy Jeanes, Tanya Mengel, Samantha Ecarnot, Karen Heinrichs, Michelle Vellenoweth, Kristin Jarvis, and Rick Fiene.

Research Team: Sonya Stevens, Alisa Hendrickson, Cindy Jeanes, Derek Pardy, Debbie Thompson, and Rick Fiene.

For additional information regarding this research validation study and report, please contact:

NARA: National Association for Regulatory Administration. http://naralicensing.org/key-indicators

REFERENCES

ACF, (2015). *Caring for Our Children: Basics*, Department of Health and Human Services, Administration for Children and Families, Washington, DC.

Fiene, (2019a). *The Saskatchewan Key Indicator System: The First Step in Developing a Differential Monitoring Approach*, National Association for Regulatory Administration.

Fiene, (2019b). *The Saskatchewan Centre and Home-Based Weighted Risk Assessment Study*, National Association for Regulatory Administration.

Fiene, (2020). *Validation Research Studies of Key Indicator and Risk Assessment Methodologies in the Province of Saskatchewan*, National Association for Regulatory Administration.

Fiene, (2022). *Regulatory Compliance Scale*, Research Institute for Key Indicators.

Fiene, (2002a). *13 Indicators of Quality Child Care: Research Update*, Office of the Assistant Secretary for Planning and Evaluation, 200 Independence Avenue S.W., Room 450G, Hubert Humphrey Building, Washington, DC 20201.

Fiene, (2014a*). Qualistar Rating Key Indicator Study*, National Association for Regulatory Administration.

Lahti, Elicker, Zellman, Fiene, (2014b). Approaches to validating childcare quality rating and improvement systems (QRIS): Results from two states with similar QRIS type designs*, Early Childhood Research Quarterly*, Special Issue.

Fiene, (2019c*). Parental Guide to Choosing Safe and Healthy Child Care*, Office of the Assistant Secretary for Planning and Evaluation, 200 Independence Avenue S.W., Room 450G, Hubert Humphrey Building, Washington, DC 20201.

Fiene, Greenberg, Bergsten, Fegley, Carl, Gibbons, (2002b). *Pennsylvania Early Childhood Program Quality Study*, Harrisburg, Pa: Governor's Office.

Fiene & Kroh, (2000). *Licensing Measurement and Systems,* part of the *NARA Licensing Curriculum*, Virginia: National Association for Regulatory Administration.

Also, check out the following websites for additional Differential Monitoring Reports: https://rikinstitute.com or https://www.naralicensing.org/key-indicators

Appendix

Saskatchewan's Early Learning and Child Care Program Quality Key Indicator Instrument (SKECPQI)

The Saskatchewan Program Quality Work Group[1]

March 2023

INTRODUCTION and BACKGROUND to SKECPQI

Ten Quality Key Indicators (QKI) make up the Saskatchewan's Early Learning and Child Care Program Quality Key Indicator Instrument (SKECPQI). The details about each of the Quality Indicators and data collection instructions in order to obtain the necessary data to determine if a program meets the Key Quality Indicators are delineated below for each quality key indicator. Part 1 - Quality Key Indicators (QKI) 1 – 5 will be collected via record or document review, interviewing individuals, or observation. Part 2 - Quality Key Indicators (QKI) 6 – 10 will be collected via observations in the classrooms throughout the assessment.

These ten quality key indicators were taken from previous studies conducted over the past 40 years by Dr Richard Fiene utilizing the Regulatory Compliance Key Indicator metric (RCKIm) that he developed in the late 1970's. These QKI have held up over time and have now been coupled together into this tool and being pilot tested in the Province of Saskatchewan. The original tool was reviewed by a Provincial Ministry of Education Work Group who met during 2019-2020 and made some revisions to the original tool. All these changes are reflected in this version of the SKECPQI (2023).

---

1) Saskatchewan Program Quality Work Group: Kim Taylor, Derek Pardy, Cindy Jeanes, Tanya Mengel, Samantha Ecarnot, Karen Heinrichs, Michelle Vellenoweth, Kristin Jarvis, and Rick Fiene.

PART 1 – Record/Document Review, Interview, Observation Quality Indicators

INDICATOR 1): Number of ECE III Educators

Assessors will review staff records in order 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 in order to come up with a percent for the center.

How to Measure:

Go to the ***Staff Information Summary*** form to obtain the data for this item. There are two particular columns that will do this. Under Certification: *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 Guide.*** The program is 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 in order 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.

NOTE: The examples are mutually exclusive and are not additive; the first example is the highest scored, the third example the least scored. After 1-3 are determined, then do the last example.

- 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:

For quality key indicators 6, 7 and 8, it is recommended that the licensing consultant refer to the appropriate Environmental Rating Scale (ERS) tool as a reference tool because these indicators are taken directly from these tools. It is also recommended that these be assessed/observed throughout the assessment and not just during key activity times. Please follow the specific instructions and examples as delineated below and in the appropriate ERS tool: ECERS (Items 12 and 13) or ITERS (Item 12). These specific instructions and examples are provided within this tool for ease of administration and data collection. If there are several preschool aged classrooms randomly select one to do your 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 conversations.

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: nontalking 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)

NOTE: If there is an infant, toddler or combined infant/toddler classroom that needs to be assessed, then use the following ITERS item directly from the ITERS Tool (Item 12), if there is not an infant toddler classroom, then skip to Indicator 8.

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 you are assessing.

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, rephrases their comments, engages 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.

10 Observations:

10.1	2	3	4	5	6	7	8	9	10.10

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***

SKECPQI 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: _____

Note:

Members of the Original Saskatchewan Program Quality Work Group are the following:

Ministry of Education: Kim Taylor, Derek Pardy, Cindy Jeanes, Tanya Mengel, Samantha Ecarnot, Karen Heinrichs, Michelle Vellenoweth, Kristin Jarvis, and NARA Consultant: Rick Fiene.

Additional Information contact: Derek Pardy, Government of Saskatchewan, Senior Policy Analyst, Early Years, Ministry of Education, 2-2220 College Ave, Regina, SK, Canada S4P 4V9.

Additional Information regarding the psychometrics of the tool contact: Richard Fiene, Ph.D., Research Psychologist, Research Institute for Key Indicators & Penn State University. RFiene@RIKInstitute.com or RFiene@NARALicensing.org

10/2020; 4/2021; 1/2023; 2/2023; 3/2023 versions

After completing your observations, reviewing all documentation, and interviewing staff, when necessary, please transfer all your results to the Summary Table below. If there was not an infant classroom, please note here, no infant classroom: _____. If there was not a toddler classroom, please note here, no toddler classroom: _____. If there was not a preschool classroom, please note here, no preschool classroom: _____.

Key Q Indicator	***Quality Indicator Content***	***Scale Source***	***Potential Score***	***Actual Score***
QKI 1	Professional Development	NAEYC	1-4	1, 2, 3, 4
QKI 2	The Environment	Saskatchewan	1-4	1, 2, 3, 4
QKI 3	Curriculum and Assessment	NAEYC	1-4	1, 2, 3, 4
QKI 4	Family Engagement I	QRIS	1-4	1, 2, 3, 4
QKI 5	Family Engagement II	QRIS	1-4	1, 2, 3, 4
QKI 6	Communication (Preschool)	ECERS	1-4 or NA	1, 2, 3, 4, +, NA
QKI 7	*Infant Classroom*	*ITERS*	*1-4 or NA*	1, 2, 3, 4, +, NA
QKI 8	Reasoning Skills (Preschool)	ECERS	1-4 or NA	1, 2, 3, 4, +, NA
QKI 9	Listen Attentively	CIS	1-4	1, 2, 3, 4
QKI 10	Speak Warmly	CIS	1-4	1, 2, 3, 4

Notes:
Use ITERS if: (Infants) (B-1yr)
Use ITERS if: (Toddlers) (1yr-2yr)
Use ECERS if: (Preschoolers) (3yr+)

SKECPQI/Infant (administer QKI items 1-5, 7, 9-10) (Scores 8-32)

SKECPQI/Toddler or Preschool (administer QKI items 1-5, 7, 9-10) (Scores 8-32) or (administer QKI items 1-6, 8-10) (Scores 9-36). Mixed age group (administer QKI items 1-10) (Scores 10-40)

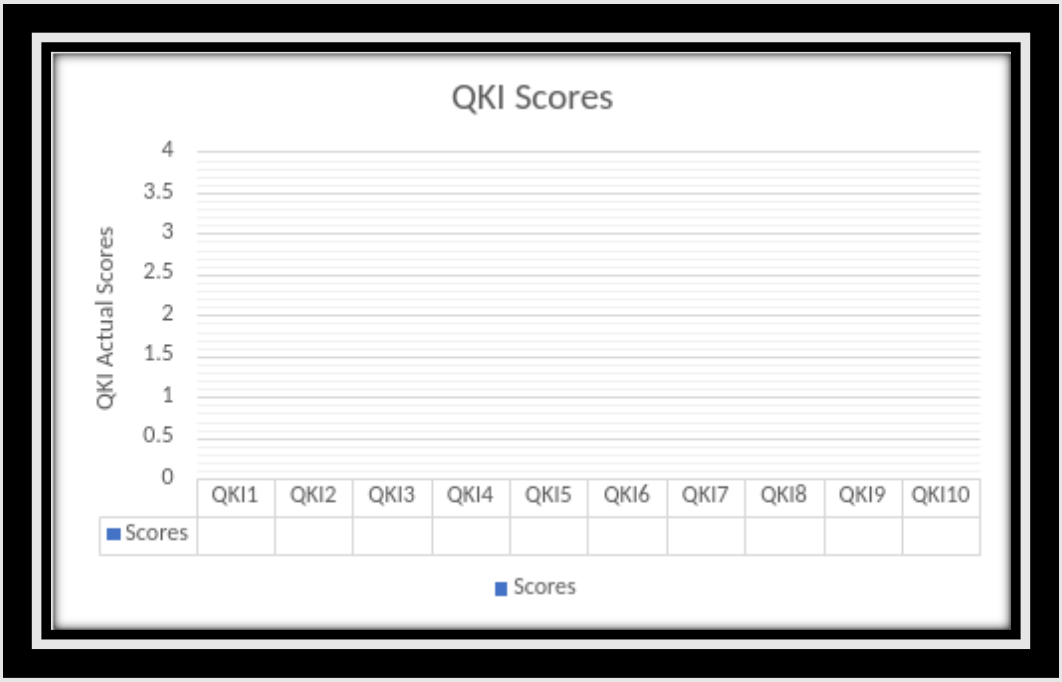
SKECPQI/Preschool (administer QKI items 1-6, 8-10) (Scores 9-36)

All the above 10 quality indicators (SKECPQI) have been taken from other sources having been identified in Quality Indicator Studies conducted by Dr Richard Fiene from 1980 – 2020. Please refer to the source documents for details on their creation: *ECERS, ITERS, QRIS/INQUIRE, CIS/Arnett, NAEYC, SASKATCHEWAN PLAY & EXPLORATION.* For additional information, reports, and publications related to these studies, please go to https://www.naralicensing.org/key-indicators Or https://rikinstitute.com/publications/

SKECPQI: SASKATCHEWAN EARLY CHILDHOOD PROGRAM QUALITY INDICATORS

CHART/GRAPH

	Scores
QKI1	
QKI2	
QKI3	
QKI4	
QKI5	
QKI6	
QKI7	
QKI8	
QKI9	
QKI10	
TOTAL	



QKI and key elements/sub items and comments Scoresheet:

QKI1 _____ 1.1 _____ 1.2 _____ Comments: _____

QKI2 _____ %

2.1 _____ Comments: _____

2.2 _____ Comments: _____

2.3 _____ Comments: _____

2.4 _____ Comments: _____

2.5 _____ Comments: _____

2.6 _____ Comments: _____

2.7 _____ Comments: _____

2.8 _____ Comments: _____

2.9 _____ Comments: _____

2.10 _____ Comments: _____

2.11 _____ Comments: _____

QKI3 _____ %

3.1 _____ Comments: _____

3.2 _____ Comments: _____

3.3 _____ Comments: _____

3.4 _____ Comments: _____

QKI4 _____ %

4.1 _____ Comments: _____

4.2 _____ Comments: _____

4.3 _____ Comments: _____

QKI5 _____ Points

5.1 _____ Comments: _____

5.2 _____ Comments: _____

5.3 _____ Comments: _____

5.4 _____ Comments: _____

QKI6 _____ Level

6.1 _____ Comments: _____

6.2 _____ Comments: _____

6.3 _____ Comments: _____

6.4 _____ Comments: _____

6.5 _____ Comments: _____

6.6 _____ Comments: _____

6.7 _____ Comments: _____

6.8 _____ Comments: _____

6.9 _____ Comments: _____

QKI7 _____ Level

7.1 _____ Comments: _____

7.2 _____ Comments: _____

7.3 _____ Comments: _____

7.4 _____ Comments: _____

7.5 _____ Comments: _____

7.6 _____ Comments: _____

7.7 _____ Comments: _____

7.8 _____ Comments: _____

7.9 _____ Comments: _____

7.10 _____ Comments: _____

7.11 _____ Comments: _____

7.12 _____ Comments: _____

QKI 8 _____ Level

8.1 _____ Comments: _____

8.2 _____ Comments: _____

8.3 _____ Comments: _____

8.4 _____ Comments: _____

8.5 _____ Comments: _____

8.6 _____ Comments: _____

8.7 _____ Comments: _____

8.8 _____ Comments: _____

QKI9 _____ Level

9.1 _____ Comments: _____

9.2 _____ Comments: _____

9.3 _____ Comments: _____

9.4 _____ Comments: _____

9.5 _____ Comments: _____

9.6 _____ Comments: _____

9.7 _____ Comments: _____

9.8 _____ Comments: _____

9.9 _____ Comments: _____

9.10 _____ Comments: _____

QKI10 _____ Level

10.1 _____ Comments: _____

10.2 _____ Comments: _____

10.3 _____ Comments: _____

10.4 _____ Comments: _____

10.5 _____ Comments: _____

10.6 _____ Comments: _____

10.7 _____ Comments: _____

10.8 _____ Comments: _____

10.9 _____ Comments: _____

10.10 _____ Comments: _____

Quality Key Indicators (QKI)	Elements/Items	Data Collection
1	1.	Record Review
2	11	Policy, Records, Interviews
3	4	Policy, Records, Interviews
4	3 . . .	Policy, Records, Interviews
5	4	Policy, Records, Interviews
6	9	Observation
7	12	Observation
8	8	Observation
9	10	Observation
10	10	Observation
TOTAL	**Potential Score = 78**	**Actual Score Obtained = _____**