RIKI/NARA

Licensing Measurement and Monitoring Systems

Regulatory science applied to human services regulatory administration

Richard Fiene, PhD
RESEARCH INSTITUTE FOR KEY INDICATORS

Suggested citation: Fiene, R. (2023). *Licensing Measurement and Monitoring Systems: Regulatory Science Applied to Regulatory Administration*. 3rd Ed. Elizabethtown, PA: Research Institute for Key Indicators (RIKILLC)

ISBN: 979 - 8 - 8370598 - 2 - 7

This eHandBook is the text to be used along with the NARA: National Association for Regulatory Administration's Licensing Measurement and Systems course which is part of the *NARA Licensing Curriculum*. This text will provide the learner with a basic introduction to licensing measurement and program monitoring state of the art key elements and principles.

Contents

RESEARCH INSTITUTE FOR KEY INDICATORS

NATIONAL ASSOCIATION FOR REGULATORY ADMINISTRATION

THE PENNSYLVANIA STATE UNIVERSITY

Preface

The reason for writing this eHandBook is to provide a short easy to read introduction to licensing measurement and monitoring systems for licensing researchers and administrators, and for regulatory scientists and policymakers. It is not intended to be a comprehensive, in-depth analysis of licensing measurement. This eHandBook provides only a basic introduction to licensing measurement. For those readers who are interested in doing a deeper dive into licensing measurement, I recommend NARA's - National Association for Regulatory Administration's *Licensing Curriculum* and their online courses offered through the University of Southern Maine.

This eHandBook will provide the basics to get the interested learner pointed in the direction of learning more about the topic. As one will see licensing measurement has its challenges and idiosyncrasies which will need to be addressed by researchers and scientists. When I started this journey 50 years ago, I was somewhat taken aback by the different data distributions I encountered in regulatory science. Back then, regulatory science was not well formulated and program monitoring related to licensing was more qualitative (case notes) rather than quantitative. But I got really interested in public policy and macro-systems which seemed to have more and more impact on children and their families. This was the beginning of governmental rule promulgation, and it was an exciting time to be on the cusp of this new research area.

I learned very quickly that I had to make several adjustments to the statistical methods I learned in graduate school to be able to analyze licensing data and measure regulatory compliance. Several of the theories and methodologies were controversial when I proposed them because they went counter to the prevailing paradigms at that point in the 1960-70's. However,

over time with many replications and validations, the new conceptual framework was accepted in the licensing and the regulatory science research literatures.

There are many people to thank over the years and obviously this has been a group effort in applying regulatory science to early care and education and then expanding it to human services and hopefully beyond. What I have found in my most recent readings is that regulatory science is being applied in many different content silos from the FDA, to economics, to banking, and of course within the human services particularly adult and child residential services. What appears to be lacking is a unifying theory that goes across these disparate content areas. That is why I think the introduction of the Regulatory Compliance Theory of Diminishing Returns is such an important contribution when we think about licensing/regulatory measurement and monitoring systems. The theory has become the foundation for the development of the methodologies and metrics presented in this eHandBook such as key indicators, risk assessment, differential monitoring, instrument-based program monitoring, integrative quality monitoring, skewed data distributions, nominal and ordinal measurement scaling, how best to deal with false negatives in decision making, and the balancing act between regulatory compliance & quality programming.

As I said earlier this eHandBook needs to be read along with the published materials on the Research Institute for Key Indicators: RIKI (https://RIKInstitute.com) and the National Association for Regulatory Administration: NARA (https://www.naralicensing.org/key-indicators) websites. It is not intended as a standalone text for licensing measurement.

Rick Fiene, PhD, Research Psychologist & Regulatory Scientist
Research Institute for Key Indicators, Penn State University, &
National Association for Regulatory Administration

March 2023

Chapter 1

Introduction to Licensing Measurement and Systems

This first chapter provides the learner with an introduction and overview to licensing measurement and systems. The ehandbook is sponsored by NARA – National Association for Regulatory Administration. NARA is the prominent international organization dealing with human services licensing. This ehandbook is part of the *NARA Licensing Curriculum* which you can find out more about by visiting NARA's website (https://www.naralicensing.org/nara-licensing-curriculum). NARA also offers a course by the same name and this book is the eTextBook for that course; it is intended to be used in conjunction with the *NARA Licensing Measurement course*.

The NARA course will provide the learner with the major tenets of licensing measurement. The learner will discover as they go through this book that measurement in licensing is very different than other measurement systems found in many of the various social and human services. It has some very unique and idiosyncratic aspects which will provide us with increasing challenges in coming up with specific metrics in determining regulatory compliance.

The field of regulatory science is a very young field. Although regulations have been kicking around for well over 100 years, the science behind regulations is probably a quarter of this time. So, there is not a great deal of empirical evidence to draw upon which is discouraging but it is very encouraging and exciting at

the same time because so much needs to be accomplished in establishing regulatory science's theory.

A great deal has been written in the past 20 years about regulatory science but there has not been a book written about measurement. It is hoped that this book will begin that discussion. It is also hoped that data driven via regulatory science will begin to inform regulatory administration and policy more clearly as we move forward.

This specific chapter will provide the conceptual framework and overview to licensing measurement and systems of regulatory compliance. It will provide the parameters of the book's organization and what will be covered throughout.

The other chapters to be covered in this book are the following:

1. ***Overview/Introduction***
2. ***Conceptual/Theoretical Framework***
3. ***Principles of Instrument Design***
4. ***Regulatory Compliance and Program Quality***
5. ***Coordinated Program Monitoring, Differential Monitoring & Integrated Monitoring***
6. ***What Research Tells Us; What Research Doesn't Tell Us***
7. ***Future Directions***

The book is organized into the above 7 chapters. The book is short and provides the basics to licensing measurement and systems. It is a quick read for regulatory scientists and regulatory policy makers as well as licensing administrators. It can be read as a standalone text although it was intended as the textbook for the *NARA Licensing Measurement* course, and it is recommended to be used with that course.

The *NARA Licensing Measurement* course is approximately 45 hours in length and is organized into anywhere from 7 - 13 classes. It is equivalent to a three-credit course offered at most institutions of higher education. Each class is organized into

the following: an overview to what will be covered in the specific class followed by annotated PowerPoint slides, followed by a series of readings to support the specific lecture/PowerPoint slides (I will be referring to these various resources throughout this text). For the learner who wants to get a thorough grounding in licensing measurement and its accompanying program monitoring systems, I highly recommend them taking the course.

This book and the course are self-paced and are geared to the individual learner. It is totally self-contained meaning that all the necessary content is contained with the thirteen classes. If a learner just wants to get an overview of what licensing measurement is all about, then reading this short ebook will be a great start. You can always check out any of the publications that are available on the RIKI Institute website (https://rikinstitute.com/publications/). However, if a learner does have a specific question related to this textbook or if they are interested in taking the course and would like to get in touch with Dr Fiene, here is his contact information to reach out (email address is the best way to contact Dr Fiene):

Dr Richard Fiene, Research Psychologist & Professor of Psychology & Human Development (ret)
Research Institute for Key Indicators (RIKI) & Penn State University
RFiene@RIKInstitute.com
Senior Research Consultant/Senior Regulatory Scientist
National Association for Regulatory Administration
RFiene@NARALicensing.org

Also, this ehandbook and the course will draw heavily from both the NARA and RIKI websites where many of the publications and research reside. Please feel free to go to https://RIKInstitute.com/blog/ to download any additional blog posts that may be of interest to you. As I said, all the research is in the public domain and follows an open science sharing arrangement.

The links for additional publications (see the references listed at the end of this book) for the NARA course are listed throughout this book or within the course format in the handouts section on the NARA website which you can download in their entirety or do it chapter by chapter. All course materials will be provided in either the lectures section or the handouts section of the class.

I thought it would be helpful to provide a bit of my background which will help the reader to put in context the content of this text and the course. This textbook is written by a research psychologist who has spent his career in improving childcare quality through an early childhood program quality indicator model of training, technical assistance, quality rating and improvement systems, professional development, mentoring, licensing, risk assessment, differential program monitoring, key indicators, and accreditation. While content wise, I spent my career in early care and education, I evolved into a regulatory scientist because of the various positions I held within governmental service and interest area that focused on public policy, macro systems, and licensing rules.

Here are some additional commentaries taken from my NIH Bio Sketch that might help to fill in some details related to my background:

Dr Fiene is a retired professor of human development & psychology (Penn State University) where he was department head and founding director of the Capital Area Early Childhood Research and Training Institute. He is presently President & Senior Research Psychologist/Regulatory Scientist for the Research Institute for Key Indicators.

Dr Fiene is regarded as a leading international researcher/scholar on human services licensing measurement and differential monitoring systems. His regulatory compliance

law of diminishing returns has altered human services regulatory science and licensing measurement dramatically in thinking about how best to monitor and assess licensing rules and regulations through targeted and abbreviated inspections.

His research has led to the following developments: identification of herding behavior of two-year old's, national early care and education quality indicators, mathematical model for determining adult child ratio compliance, solution to the trilemma (quality, affordability, and accessibility) in child care delivery services, Stepping Stones to Caring for Our Children, online coaching as a targeted and individualized learning platform, validation framework for early childhood licensing systems and quality rating & improvement systems, an Early Childhood Program Quality Improvement & Indicator Model, Caring for Our Children Basics, and has led to the development of statistical techniques for dealing with highly skewed, non-parametric data distributions in human services licensing systems.

Organization of the eHandBook

In the following chapters, the reader will find how licensing measurement is very unique when it comes to regulatory science. By knowing these unique characteristics, it will be easier to administer and monitor programs governed by licensing rules. This can be of benefit to those who administer licensing agencies and to those who are asked to conduct research on regulatory policies and compliance.

In Chapter 1, this chapter, provides the basic introduction and overview to the ehandbook as well as a short history of licensing measurement and a timeline for early care and education standards focusing on *Caring for Our Children: Basics*.

In Chapter 2, the reader will be introduced to the theoretical and conceptual foundations to licensing measurement. The regulatory compliance theory of diminishing returns will be introduced which has had a tremendous impact on human services regulatory science and administration. Paradigm alternatives will be suggested that have guided regulatory science over the past several decades.

In Chapter 3, the principles of instrument design will be addressed. Obviously without a reliable and valid system of measurement it will be simply garbage in and garbage out. This is an area when it comes to instrument design that gets the short end of the stick many times. The level of measurement will be addressed and its impact on the types of statistics selected. Also, how best to design data bases will be addressed.

In Chapter 4, regulatory compliance and program quality will be discussed presenting it on a dichotomous ten-point polemic for the regulatory science field to consider. This is a very important chapter building off the theory of regulatory compliance of diminishing returns introduced in chapter 2.

In Chapter 5, the essence of program monitoring systems is introduced along with differential monitoring and its two major methodologies of key indicators and risk assessment. This chapter gets us thinking about what a licensing measurement will look like administratively.

Chapter 6 deals with the research literature what we know and what still needs to be addressed, the gaps in our regulatory science knowledge base. Examples are provided of success stories across the USA and internationally.

The last chapter, Chapter 7, provides us with where do we go from here with licensing measurement and systems. What are the next steps.

There are a series of technical research notes that form an appendix which help to clarify the body of the text and add some details. This is followed by a comprehensive reference listing and then several figures, charts, graphs, and displays which depict various concepts presented in the text.

A Brief History of Licensing Measurement

The history of licensing measurement and regulatory compliance has actually a rather long lineage but is still in its infancy in terms of development. In the early stages most licensing visits and inspection results were recorded via anecdotal records/case records with the licensing staff recording their results in more social work note taking. It was a qualitative type of measurement with very little quantitative measurement occurring with the exception of basic demographics, number of clients, number of caregiving staff, etc… This qualitative approach worked very well when there were not many programs to be monitored and there were sufficient licensing staff to do the monitoring and conduct the inspections.

This all started to change in the 1980's when Instrument Based Program Monitoring (IPM) was introduced and started to be adopted by state licensing agencies throughout the United States. Just as a footnote, this brief history is pertinent to the USA and does not include other countries although the Canadian Provinces have followed a similar route as the USA. The reason for the introduction of an IPM approach was the tremendous increase in early care and education programs in the 1960's and 1970's. It was difficult for licensing staff to keep up with the increased number of programs in their monitoring efforts. There needed to be a more effective and efficient methodology to be employed to deal with these increases.

A very influential paper was written in 1985 and published in *Child Care Quarterly* which introduced IPM along with Licensing Key Indicators, Risk Assessment (Weighting), and

Differential Monitoring (Abbreviated Inspections). This paper outlined the various methodologies and their use by a consortium of states to test the viability of this new approach to licensing measurement, regulatory compliance, and program monitoring. Also, the terminology has changed over the decades. Back in 1985 weighting was used rather than risk, abbreviated inspections were used rather than differential monitoring, targeted monitoring, or inferential monitoring. All these terms can be used interchangeably as they have been over the years, but the first introduction of them back in 1985 utilized weighting and abbreviated inspections.

In the early 1990's the risk assessment methodology was used to develop *Stepping Stones to Caring for Our Children*, the comprehensive national health and safety standards for early care and education (ECE) programs in the USA. This was a major development in attempting to develop national voluntary standards for child care in the USA.

It was during this time that two other very significant discoveries occurred related to licensing data distributions: 1) Licensing data are extremely skewed and do not follow a normal curve distribution. This fact has a significant impact on the statistics that can be used with the data distributions and how data analyses are performed. For example, data dichotomization is warranted with licensing data; 2) Regulatory compliance data are not linear when compared to program quality measures but are more plateaued at the substantial and full regulatory compliance levels. The data appear to follow the Law of Diminishing Returns as compliance moves from substantial to full (100%) regulatory compliance. This finding has been replicated in several studies and has been controversial because it has led to the issuing of licenses to programs with less than full compliance with all rules/regulations/standards. These two discoveries have been very influential in tracking developments in licensing measurement since their discoveries.

In the new century as states began to adopt the various methodologies it became necessary to have a standardized approach to designing and implementing them. The National Association for Regulatory Administration (NARA) took up this role and in 2000 produced a chapter on Licensing Measurement and Systems which helped to guide states/provinces in the valid and reliable means for designing and implementing these methodologies. In 2002 a very important study was conducted by the Assistant Secretary's Office for Planning and Evaluation (ASPE) in which they published the *Thirteen Indicators of Quality Health and Safety and a Parent's Guide* to go along with the research. This publication further helped states as they revised their licensing and program monitoring systems for doing inspections of early care and education facilities based upon the specific indicators identified in this publication. Both publications have been distributed widely throughout the licensing world.

During the first decade of the new century, *Stepping Stones for Caring for Our Children* went through a second edition. This publication and the ASPE publications were very useful to states as they prepared their Child Care Development Fund (CCDF) plans based upon Child Care Development Block Grant (CCDBG) funding.

From 2010 to the present, there have been many major events that have helped to shape licensing measurements for the future. *Caring for Our Children Basics (CFOCB)* was published and immediately became the default voluntary early care and education standards for the ECE field. The *CFOCB* is a combination of the risk assessment and key indicator methodologies. Three major publications by the following Federal agencies: HHS/ACF/USDA: Department of Health and Human Services/Administration for Children and Families/United States Department of Agriculture, OCC: Office of Child Care, and ASPE: Assistant Secretary's Office for Planning and Evaluation dealing with licensing and program monitoring strategies were published. These publications will

guide the field of licensing measurement for years to come. The Office of Head Start developed and implemented their own Head Start Key Indicator (HSKI) methodology. And in 2016, CCDBG was reauthorized and differential monitoring was included in the legislation being recommended as an approach for states to consider.

Most recently, the Office of Head Start is revising their monitoring system that provides a balance between compliance and performance. This system revision will go a long way to enhancing the balance between regulatory compliance and program quality. Also, there has been experimentation with an *Early Childhood Program Quality Indicator* instrument combining licensing and quality indicators into a single tool. These two developments help with breaking down the silo approach to measurement where licensing and quality initiatives are administered through separate and distinct approaches such as licensing versus professional development systems versus quality rating and improvement systems. A paradigm shift in which an *Early Childhood Program Quality Improvement and Indicator Model* is proposed. The paradigm shift should help to make licensing measurement more integrated with other quality initiatives.

The licensing field continues to make refinements to its measurement strategies in building a national/international regulatory compliance data base. More and more is being learned about the nuances and idiosyncrasies of licensing data, such as moving from a nominal to an ordinal driven data system. For example, NARA and the Research Institute of Key Indicators (RIKI) have entered into an exclusive agreement for the future development of licensing measurement strategies via differential monitoring, key indicators for licensing and program quality, and risk assessment approaches. Several validation studies have been completed in testing whether the various methodologies work as intended. A significant Office of Program Research and Evaluation (OPRE) *Research Brief* which developed a framework for conducting validation studies

for quality rating and improvement systems has been adapted to be used in licensing measurement.

For additional updates to licensing measurement, please check out and follow these RIKINotes Blog posts. There are and will be many examples of licensing measurement enhancements. Also, although much of the research on licensing measurement has been completed in the ECE field, the methodologies, models, systems, and approaches can be utilized in any human service arena, such as child residential or adult residential services. Also, NARA's chapter in their Licensing Curriculum has been developed into a full-blown course, please go to the following web page for additional information (https://www.naralicensing.org/key-indicators).

A Timeline of ECE Standards and Program Monitoring in the United States: *Caring for Our Children Basics—probably the best example of a health and safety standards tool*

It all started in and around 1965 when the Federal government got into early care and education (ECE) in earnest with Head Start and federally funded day care for low-income families. It started off slowly but began to pick up momentum with exciting studies and research applying principles from developmental psychology to policy making. Researchers and policy makers wanted to make sure that these new programs were not detrimental to young children since our frame of reference were children being raised in orphanages and the ultimate outcome for children was not positive. Would ECE have the same impact?

Issues around quality, appropriateness of standards, and demonstration programs became the focal point of federal research funding. The focal point of this essay is on the appropriateness of the ECE standards and the resulting monitoring systems that were to become key to the federal involvement in early care and education. This essay will be

organized by the following 50 years neatly broken out by each decade to get us from this beginning in 1965 until the publication of *Caring for Our Children Basics* in 2015 by the federal government, the Administration for Children and Families, U.S. Department of Health and Human Services. A look at the 2020 decade with a future note is also appended to this essay.

1970s

During the 1970's, the federal government became concerned about what were to be the standards for this new national program related to federally funded ECE for low-income families and their children. Head Start was a separate entity and we will revisit Head Start later but our focus for now is on the federally funded programs which became known back then generically as day care. This nomenclature changed to child care and to finally early care and education (ECE) during this 50-year history. The initial standards for day care were the *Federal Interagency Day Care Requirements (FIDCR)*. A very large appropriateness research study led by Abt Associates to determine what were the most salient standards and their intended impact on children while in day care was conducted during this decade. These standards were to be federally mandated requirements for any program receiving federal funding. This is where group size and adult-child ratios standards became such important safeguards and surrogates for children's health and safety in day care programs.

It also became of interest for the federal government to design the monitoring system that would determine compliance with the FIDCRs. But it became clear to the original designers of this new system that the monitoring of the FIDCR was going to be difficult to do across the full USA. So, the question became, is there a way to monitor the standards in the most effective and efficient manner? This question and the future of the FIDCR were to be altered and put on hold once we moved into the next decade.

1980s

A change in federal administration and a resulting change in philosophy related to the federal role in America altered many things and one of them was the relationship of the federal government and the states. Rather than the federal government mandating day care requirements, the focus changed with the locus of control moving from the federal level to the state level via block grant funding with very few federal requirements. This meant a moratorium to FIDCR and its ultimate demise. The federal government was not going to be in the business of providing day care, this was going to be the jurisdiction of the states. Head Start did become the exception to this rule with its own standards and monitoring system.

The focus of federal funding switched from the national to the state level in determining compliance with each state's respective child care licensing rules and not with an overarching FIDCR. There was still interest in making these state monitoring systems as effective and efficient but there was no interest in the federal government determining what these requirements would be. Two monitoring approaches grew out of this need for effectiveness and efficiency: risk assessment and key indicators. These two approaches were originally designed and implemented as part of a federally funded project called the Children's Services Monitoring Transfer Consortium in which a group of five states: New York, Michigan, Pennsylvania, West Virginia, and California teamed up to explore their most effective and efficient monitoring systems and begin transferring these systems to one another and beyond.

These two monitoring approaches were tested in the above respective states and it was determined that their impact had a positive effect on the children who were in those day care centers. This was a major finding, similar to the FIDCR appropriateness study, in which these approaches provided safeguards related to the health and safety of children while in day care.

1990s

By the 1990s, it became clear that the federal government had pretty much drawn back from any leadership role in having mandated federal requirements when it came to health and safety in child care. It was left to national ECE advocates who were positioned within the federal government (Administration for Children and Families; Maternal and Child Health Bureau) as well as throughout the USA with national and state agencies and organizations (American Academy of Pediatrics; American Public Health Association, National Resource Center for Health and Safety in Child Care) that saw a need for child care health and safety recommendations at least. If we could not have requirements, we could at least have recommendations and provide guidance to child care programs throughout the USA.

This led to the first edition of *Caring for Our Children* which was a comprehensive set of child care health and safety standards. It was a major game changer for the ECE field because now there was a universal set of standards based upon the latest research literature for states to use as they considered revising and updating their respective state licensing child care rules.

But there was a problem. *Caring for Our Children* was a comprehensive set of health and safety standards which was their strength but at the same time it was their weakness. They were so comprehensive (well over 500 well researched standards) that they were intimidating and it was difficult to determine where to begin for the states.

Several researchers remembered the two approaches to monitoring designed in the previous decade and wondered if they could be helpful in focusing or targeting which of the standards were the most critical/salient standards. The risk assessment approach to monitoring appeared to have the most immediate applicability and *Stepping Stones* to *Caring for Our Children* was born. This document clearly articulated which of the 500+ *Caring for Our Children* standards placed children at

greatest risk for mortality or morbidity by not being in compliance with the respective standard. Since the early 1990s, *Caring for Our Children* and *Stepping Stones to Caring for Our Children* have gone through three editions and have become very important resources to state licensing agencies as they revise, update and improve their ECE rules.

2000s

In this decade several federal and national organizations began to use *Caring for Our Children* standards in innovative ways to measure how well ECE looked at a national level. The Assistant Secretary's Office for Planning and Evaluation in the U.S. Department of Health and Human Services published the *Thirteen Indicators of Quality Child Care* based upon a core set of predictor standards from *Caring for Our Children*. These were standards that predicted overall compliance with all the standards and were seen as an efficient monitoring system. NACCRRA (National Association for Child Care Resource and Referral Agencies) began publishing a national report card on how well states met specific standards and monitoring protocols based upon similar predictor standards from *Caring for Our Children*.

These efforts helped states to make significant changes in their ECE rules in their respective states and in a very voluntary way suggested a means for national standards for the ECE field although we would need to wait until the next decade in order to see such a published document of national ECE health and safety standards for early care and education: *Caring for Our Children Basics*.

2010s

By the 2010s, ECE had grown into a very large but unwieldly assortment of programs with varying levels of quality. Again because of major federal funding, the Child Care Development Block Grant, along with changes and enhancements in professional development, accreditation systems, quality rating

and improvement systems, the ECE landscape had become more complex and less easy to navigate. And rather than coming together it was clearly more fragmented than ever.

We had very minimal requirements for the federal funding and most of these requirements were geared to the state agency using the state's respective licensing rules as the threshold for standards. This approach worked well with states with excellent licensing rules, but it wasn't working as well with states who did not have equally excellent licensing rules. We still did not have a core set of standards for ECE programs. Enter *Caring for Our Children Basics* which took the best aspects of the above two monitoring approaches, risk assessment and key indicators and molded it into this new document. This work was led by the federal government's Administration for Children and Families, U.S. Department of Health and Human Services and although the standards are still recommendations and guidance, it is our best attempt at having national standards for early care and education. It is an attempt to provide guidance to the full ECE field, child care, Head Start, preschool, and center based as well as home-based care. It would be nice to have *Caring for Our Children Basics* as the health and safety foundation for early care and education throughout the USA. I don't see this happening in my lifetime.

2020s: Looking to the Future

As a footnote to this essay, the new decade has been dealt with a major curve ball with COVID19 rearing its ugly head and ECE has been impacted greatly because of this pandemic. As of this writing we are nowhere closer to a solution to getting ECE programs back on line. If anything, the pandemic really demonstrated the fragility of the ECE system we have built over the past 50 years and it clearly has not done very well. My hope is that we can learn from the past 50 years and not continue another 50 years along the same route; although I am guessing that many ECE advocates would be glad to have what we had

before the pandemic because what we have right non-sustainable. We know a lot more today than what we knew back in 1965 when we were worried about would day care hurt children's development. We know today that quality ECE benefits children but unfortunately, we are no closer to attaining this today than we were 50 years ago.

Two programs that have been very successful in avoiding these pitfalls are Head Start and the national Military Child Care program. Both programs are exemplary examples of quality early care and education being provided with separate funding streams and standards. Interesting enough when the Administration for Children and Families published *Caring for Our Children Basics*, both these programs were part of the reach of the published standards. As we re-invent and re-structure ECE we should be looking to both these very successful programs for guidance.

The above history and timeline have been drawn from the early care and education field but very honestly that is where the majority of the research and development has been over the past 50 years. There have been some excursions into other human services but they don't have the details as with early care and education.

Chapter 2

Conceptual and Theoretical Underpinnings, Program Monitoring Paradigms

This second chapter provides the learner with the key conceptual and theoretical foundations related to licensing measurement. As you have seen from the first chapter and will continue to see throughout this book, licensing measurement does have some idiosyncrasies which are not present in other data distributions.

Well, the same thing can be said when it comes to the conceptual and theoretical underpinnings. One of the first limitations that will be noted is the regulatory compliance theory of diminishing returns which has tremendous implications when implementing and enforcing rules. It had always been assumed that full 100% regulatory compliance with rules was what made a high-quality program. However, in the late 1970's and into the early 1980's, it became clear that this was not the case. When this hypothesis was tested it became clear that moving from low regulatory compliance to substantial regulatory compliance did demonstrate that program quality differed significantly in the substantial regulatory compliant programs being of a higher quality than those of lower regulatory compliance. However, when one moved from the substantial regulatory compliance level to the full 100% regulatory compliance level, there was a definite plateauing effect in which the programs were not increasing in quality as previously and in some cases, actually decreased in quality.

This above result was surprising and very controversial when it was first published in the mid 1980's. Many, if not most, regulatory compliance specialists did not agree with the finding. However, this relationship has held up in many other studies

conducted since then and in other human service areas. It became the new rule in clearly demonstrating if not a decline, always a plateauing effect in moving from substantial to full compliance. Today because of all these supporting studies, the result is generally accepted and has influenced public regulatory compliance policy formulation throughout the world.

This regulatory compliance theory of diminishing returns has had tremendous impacts in how we have come to measure regulatory compliance in the licensing field. Rather than viewing it in a linear modality, it suggested that a more targeted, non-linear modality or metric might be more effective and efficient. Rather than focusing on full regulatory compliance it suggested that a key indicator, abbreviated, or targeted monitoring of rules was a better approach.

Without the regulatory compliance theory of diminishing returns, the focus on what has become differential monitoring or targeted monitoring would never have occurred. There would have been no need to move from always requiring full 100% regulatory compliance with all rules. This is a very important distinction and you, the learner, will see many applications and implications as you move through the chapters in this text.

Moving from the Theory to the Conceptual

Conceptually, licensing measurement is built around obviously licensing but there are other systems which impact on licensing which are demonstrated in the first licensing measurement class when one compares the various regulatory and non-regulatory systems in the *Morgan Model - Methods for Achieving Quality Child Care*. There are contractual systems, such as QRIS (Quality Rating and Improvement System) or other types of quality initiatives. These other types of quality initiatives are non-contractual systems, such as professional development or

training or technical assistance systems; or accreditation systems.

These above systems can be integrated into a unified model called the ***Early Childhood Program Quality Improvement/Indicator Model or Differential Monitoring Logic Model and Algorithm (ECPQI2M/DMLMA)*** which is depicted in the lecture slides from the NARA Licensing Measurement Course that accompanies this text if you desire to utilize those resources and is detailed in several of the handouts. Since this will become the unifying framework when discussing licensing measurement, I would suggest that you as the learner spend a good deal of time reviewing those slides and handouts. I would think that you will want to return to them as you move through the upcoming chapters and classes as part of the NARA course to make certain you continue to understand how all the disparate pieces fit together into a uniformed whole.

By using the ***ECPQI2M/DMLMA*** (also see chapter 5 which provides a more detailed step by step guide for the development within a state licensing agency) it offers all the key elements to building an effective and efficient program monitoring system by integrating regulatory compliance and program quality and professional development systems along with differential monitoring's risk assessment and key indicator methodologies.

There are readings related to professional development that are important components to making sure that the ECPQI2M is working as it should. One of the consistent key indicators deals with professional development/training. There are examples of creative and innovative ways the training can be delivered over the internet. Pay particular attention to the iLookOut program, especially to its delivery system. Check out the https://RIKInstitute.com/publications/ website for these publications and reports, there are several articles that describe the program as well as its innovative cognitive mapping and online delivery platform.

Program Monitoring Paradigms

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

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

Rules are everywhere. They are part of the human services landscape, economics, banking, sports, religion, transportation, housing, etc... Wherever one looks we are governed by rules in one form or another. The key is determining an effective and efficient modality for negotiating the path of least resistance in complying with a given set of rules2. It is never about more or less rules, it is about which rules are really productive and which are not. Too many rules stifle creativity, but too few rules lead to chaos. Determining the balance of rules is the goal and solution of any regulatory science paradigm.

Differential/Relative versus Absolute/Full Regulatory Compliance Paradigms: this is an important key organizational element in how rules are viewed when it comes to compliance. For example, in an absolute/full approach to regulatory compliance either a rule is in full compliance or not in full

compliance. There is no middle ground. It is black or white, no shades of gray as are the cases in a differential/relative paradigm. It is 100% or zero. In defining and viewing these two paradigms, this dichotomy is the organizational key element for this paper. In a differential/relative regulatory compliance paradigm full compliance is not required and emphasis on substantial regulatory compliance becomes the norm.

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

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

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

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

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

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

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

6) Formative versus Summative: differential/relative regulatory compliance monitoring systems are formative in nature where there is an emphasis on constant quality improvement and getting better. In absolute/full regulatory compliance monitoring systems, the emphasis is on being the gate-keeper (more about the gate-keeper function in the next section on regulatory compliance/licensing and program quality) and making sure that decisions can be made to either grant or deny a license to operate. It is about keeping non-optimal programs from operating.

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

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

measurement is dealt with in the next section on regulatory compliance/licensing and program quality).

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

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

As the regulatory science and administrative fields in general continue to think about the appropriate monitoring systems to be designed and implemented, the above structure should help in thinking through what these measurement systems' key elements should be. Both paradigms are important, in particular contexts, but a proper balance between the two is probably the best approach in designing regulatory compliance monitoring systems.

Chapter 3

Instrument Design, Reliability and Validity, Statistical Methods and Databases

This third chapter provides the learner with the key principles of instrument design as it relates to licensing measurement. As you have seen there are idiosyncrasies' conceptually and theoretically and there are limitations as well, when it comes to instrument design. A major limitation with licensing data is that it is basically, nominal in nature. It fits the format of Yes or No responses. It is not ordinal in any fashion, or at least it hasn't been for the past 50 years. In fact, it is only in the past 30 or so years that licensing data moved from being predominantly qualitative to quantitative. This change started in the 1980's with the publication of Instrument based program monitoring. Prior to that most licensing studies were written as social work case studies with a great deal of narrative detail but short on data utilization that could be used at the macro level.

Instrument based program monitoring has its critics who are not overly excited about its checklist type approach. However, if a state is going to track where there are specific issues related to regulatory compliance it will be difficult unless an instrument/tool/checklist is not used in data collection. If there is continued reliance on narrative reports solely it will be difficult if not impossible to find any real patterns in the data. It is possible with the latest developments in qualitative analyses but it is not recommended as the sole means for tracking regulatory compliance. I prefer a mixed methods approach which focuses on the strengths from both the quantitative and qualitative and combines both together.

Without an instrument-based program monitoring approach it would be impossible to utilize the risk assessment and

especially the licensing key indicator predictor methodologies. In fact, it is really a pre-requisite for designing and implementing a targeted monitoring or differential monitoring approach.

In instrument design it is important to utilize the triangulation measurement strategy that looks for observation first, followed by record/document review, and then lastly by doing interviews of staff or parents. The majority of data collection should be through observations made in the classroom or facility. When observations cannot be made look for policies, files, documents that contain the necessary data and then lastly do interviews.

Reliability and Validity

This section provides the learner with the key principles of reliability and validity which are the mainstay of any measurement system. Without these two key principles we do not have a measurement system we can rely on. Reliability deals with consistency across inspectors to make certain that what is to be measured is measured accurately. Validity demonstrates that the system is working as it is supposed to. The results are what should be expected from a licensing or regulatory compliance system.

The readings and handouts provide many examples of validation studies conducted in the past decade demonstrating the validity and reliability of the licensing key indicator predictor and risk assessment methodologies (State of Washington and the Province of Saskatchewan are the best examples of these validation studies).

Since the large influx in the use of these methodologies over the past couple of decades it was incumbent upon us to determine if these methodologies were both reliable and valid. Based upon these validation studies, it can now be said with a great deal of certainty that the methodologies do what they were intended to

do. They statistically predict overall regulatory compliance and they focus on those rules that place children in greatest risk of morbidity or mortality keeping them safe. So, the tenet, which will be emphasized throughout this course "Do No Harm" is upheld!

The NARA Licensing Measurement course provide the lecture slides where an overview and the key elements to doing validation studies, while the readings and handouts provide more of the details and the results from these studies. (https://www.naralicensing.org/key-indicators)

Statistical Methods and Data Bases

This section deals with the statistical methods used and the construction of the databases in licensing. As I have said repeatedly in my writings over the years there are many limitations related to licensing measurement. The statistical methods that can be used with licensing data are limited also, because we are dealing with nominal data that are severely skewed. Non-parametric statistics is warranted and to deal with the severely skewed data, dichotomization of the data base is warranted.

Dealing with data that are not normally distributed poses some real challenges in analyzing licensing data sets. It is paramount that one runs basic descriptive statistics in assessing the mean, standard deviation, variance, skewness, and kurtosis. It will help in identifying how badly the data has outliers in a very quantitative manner. It will also help in determining where the cut scores or thresholds should be for defining the high regulatory compliance and the low regulatory compliance groups. The Fiene Licensing Predictor Rules and their respective Fiene Coefficients are determined by using the phi coefficient in determining correlations between each rule and the high/low groups for regulatory compliance. This is a statistic used with nominal data and is used a great deal in the

tests and measurement research literature invalidating testing procedures.

The databases should be saved in .csv formatting from an Excel file. It is easier to import a .csv file into SPSS or PSPP which is the preferred statistical package for conducting these analyses. But definitely any statistical package can be used as well, such as SAS for example. Outside of generating Fiene Coefficients, there are no other statistical techniques that are needed in analyzing the database.

The readings list (https://RIKInstitute.com/publications/) provides most, if not all, of the technical research notes generated by the Research Institute for Key Indicators. These tech research notes provide the latest and most up to date information about any changes in the methodologies for generating licensing key indicator predictor rules and risk assessment rules. These technical research notes are really intended for the serious licensing research and regulatory scientist to delve into. They provide the specifics to the various statistical methodologies with specific algorithms and logic modeling.

But it still important to address some of the specific statistical formulae pertinent to licensing and regulatory science data. For example, not all statistics will be relevant to licensing data because of its measurement limitations. Licensing data are nominal in nature with some instances of ordinal measurement. And there are other significant considerations, such as the skewness of the data distributions in most licensing data, non-linear nature of the regulatory compliance data when compared to quality data. So let's start with the most pertinent statistical formulae to be addressed when analyzing licensing and regulatory compliance data.

The first statistic we need to address is that of the skewness of the data. The below formula provides the basic formula for determining the skewness of the data distribution.

$$G_1 = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s} \right)^3$$

This formula is called the adjusted Fisher-Pearson standardized moment coefficient and is generally used behind the scenes in most software.

Another formula you will encounter with licensing data will be for determining the kurtosis. The following formula provides the basic formula for determining the kurtosis of the licensing data distributions.

$$Kurtosis = n * \Sigma ni\ (Yi - \bar{Y})\ 4\ /\ (\Sigma ni\ (Yi - \bar{Y})\ 2)\ 2$$

Another statistic that will be of importance is the variance of the data distribution. I have included both the population and sample variance formulae because in some cases we need to draw a sample and in other cases we have the population data. Variance is important with licensing data because it is very lacking when you really examine the data distributions which is not a good thing from a statistical point of view.

Population Variance	Sample Variance
$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$	$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$
σ^2 = population variance	s^2 = sample variance
x_i = value of i^{th} element	x_i = value of i^{th} element
μ = population mean	$\bar{x}$ = sample mean
N = population size	n = sample size

One of the most encountered statistic is the "mean" which is the average of the data distribution. Unfortunately with most licensing data the mean is not as meaningful as the "median" which is the mid point of the data distribution. I have included both because the mean is so predominant and the median and quartiles are more prevalent in licensing research. The reason for using the median over the mean is that the licensing data distributions are so severely skewed.

But the mean formula is provided below:

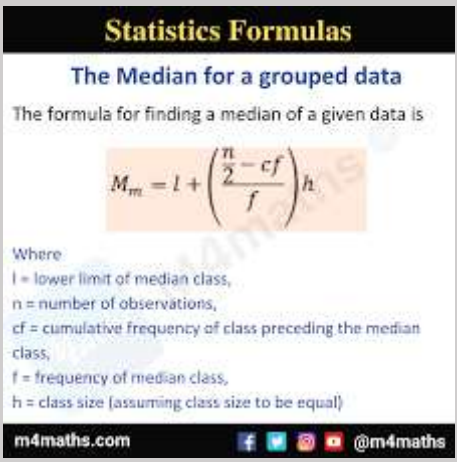
Mean of Grouped Data:

$$\bar{x} = \frac{\sum fx}{n}$$

where: $\bar{x}$ = mean
f = frequency of each class
x = mid-interval value of each class
n = total frequency
$\sum fx$ = sum of the producst of mid – interval values and their corresponding frequency

In the majority of cases, the following formula for the median will be used as a better measure of central tendency and the average score for licensing data. Also, the data are nominal in measurement which means we will be collecting frequency data, the data are not continuous, they are very discrete. Either a rule or regulation is in compliance or out of compliance. There are no metrics in between these two extremes. Data will be organized and displayed in frequency tables or cross-tabs.

Formula for calculating the median for licensing data:



Also, keep in mind that the types of analyses you will be able to accomplish will be limited because of the nature of the data measurement. When it comes to looking at relationships between data sets you will be limited to cross-tabulations and the use of the chi-square statistic. I have provided the chi-square statistic below to be used with licensing data. Other statistics which require a normal distribution or a continuous distribution cannot and should not be used.

A methodology that has been successful with licensing data has been the dichotomization of the data distribution because of the nominal measurement of compliance vs non-compliance. Generally dichotomization of data is not recommended nor warranted but I have found that this approach is very successful with licensing data distributions.

The formula for Chi-Square:

The Formula for Chi Square Is

$$\chi_c^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

where:

c = degrees of freedom

O = observed value(s)

E = expected value(s)

As one can see these statistics provide a basic grounding from an analytical point of view but it is limited because of the real limitations in the licensing data measurement characteristics.

The formula for Key Indicators is in the last section of this book within the Graphs, Charts, Figures, and Display Section.

Chapter 4

Regulatory Compliance and Program Quality

This fourth chapter provides the learner with the similarities and differences between regulatory compliance and program quality. In the second chapter the regulatory compliance theory of diminishing returns was presented which demonstrated a non-linear relationship between regulatory compliance and program quality. In this chapter, additional concepts will be presented to deal with this dynamic tension between regulatory compliance and program quality and how we can build one upon the other.

In fact, the future of licensing and regulatory compliance will be heavily influenced by this relationship between regulatory compliance and program quality. Many jurisdictions are attempting to build in quality to their rules/regulations. They are being very creative in either building separate systems (Quality Rating and Improvement Systems: QRIS) or attempting to build them right into the rules themselves in more of an ordinal format.

QRIS: Quality Rating and Improvement Systems and other Quality Initiatives

This section provides the learner with key examples from the program quality arena, such as QRIS and professional development. The ECPQI2M model presented here has these two systems prominently displayed along with the regulatory compliance or licensing system. Together they form the solid foundation for providing a very effective delivery system of services. When these are combined with risk assessment and the key indicator methodologies one can add efficiency to the effectiveness side of the equation. The next chapter, Chapter 5,

will get into more details about how to design an ECPQI2M model along with its associated logic model (DMLMA: Differential Monitoring Logic Model and Algorithm).

As was mentioned in the previous section, there is a delicate balance between regulatory compliance and program quality. At all times, the ECPQI2M is to keep both regulatory compliance and program quality in balance, to keep health & safety and quality on an even keel; but as we have seen and will see later in this course, this balancing act can get out of kilter at times.

One of the publications produced for OPRE about QRIS Validation is directly applicable to licensing measurement and has been used within this context in the validation studies that will be described in this course. This is an important application of this new framework when it comes to validation. It is not just for QRIS but can be applied to licensing as well. The state of Washington has probably some of the best examples. Please check out these resources and readings later on the RIKI website for additional examples. Look specifically for the Validation Framework Research Brief (Zellman & Fiene, 2012) published by OPRE and the state of Washington Research Agenda Report (Stevens & Fiene, 2015).

Regulatory Compliance/Licensing and Quality

This section of the chapter will delineate the differences between regulatory compliance and quality. It will provide the essential principles and elements that clearly demonstrate the differences and their potential impact on program monitoring. Obviously, there is some overlap between this section and the above section dealing with regulatory compliance monitoring paradigms. When we think about regulatory compliance measurement, we are discussing licensing systems. When we think about quality, we are discussing Quality Rating and Improvement Systems (QRIS), accreditation, professional

development, or one of the myriad quality assessment tools, such as the Classroom Assessment Scoring System (CLASS) or Environment Rating Scales (ERS's). All these systems have been designed to help improve the health and safety of programs (licensing) to building more environmental quality (ERS), positive interactions amongst teachers and children (CLASS), enhancing quality standards (QRIS, accreditation), or enhancing teacher skills (professional development).

There are ten basic principles or elements to be presented (they are presented in a binary fashion demonstrating differences): 1) "Do no harm" versus "Do good". 2) Closed system versus Open system. 3) Rules versus Indicators. 4) Nominal versus Ordinal measurement. 5) Full versus Partial compliance. 6) Ceiling effect versus No Ceiling effect. 7) Gatekeeper versus Enabler. 8) Risk versus Performance. 9) Structural versus Process Quality. 10) Hard versus Soft Data.

1) Let's start with the first principal element building off what was discussed in Chapter 2, "Do No Harm" versus "Do Good". In licensing, the philosophy is to do no harm, its emphasis is on prevention, to reduce risk to children in a particular setting. There is a good deal of emphasis on health and safety and not so much on developmentally appropriate programming. In the quality systems, such as QRIS, accreditation, professional development, Environmental Rating Scales, CLASS, the philosophy is to do good, its emphasis is looking at all the positive aspects of a setting. There is a good deal of emphasis on improving the programming that the children are exposed to or increasing the skill set of teachers, or improving the overall environment or interaction that children are exposed to.

2) Closed system versus Open system. Licensing is basically a closed system. It has an upper limit with full compliance (100%) with all rules. The goal is to have all programs fully comply with all rules. However, the value of this assumption has been challenged over the years with the introduction of the

Regulatory Compliance Theory of Diminishing Returns. With quality systems, they have a tendency to be more open and far reaching where attaining a perfect score is very difficult to come by. The majority of programs are more normally distributed where with licensing rules the majority of programs are skewed positively in either substantial or full compliance. It is far more difficult to distinguish between the really best programs and the mediocre programs within licensing but more successful in quality systems.

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

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

5) Full or None versus Gradients or Gray Area. Building off of the fourth element, licensing scoring is either full or not. As

suggested in the above elements, there is no in-between category, no gradient or gray area. This is definitely not the case with quality systems in which there are gradients and substantial gray areas. Each best practice can be measured on a Likert scale with subtle gradients in improving the overall practice.

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

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

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

9) Structural Quality versus Process Quality: when we think of structural quality, we generally think of things we can count easily, such as the number of children or teachers present in a classroom or the number of smoke alarms, etc. These are items that form the basis of rules within a licensing system. However, when we think of process quality, we generally think of things that are not as ease to measure, such as interactions between

teachers and children that are warm and engaging. This is much more difficult to measure and generally not part of licensing systems but rather program quality tools, such as the ERS and CLASS tools.

10) Hard Data versus Soft Data: this dichotomy is similar to number 9 structure quality versus process quality but adds a small dimension not present in number 9. It deals with the ease with which legal counsel can defend a specific rule or standard in a court of law. Hard rules or standards are easy to measure while soft rules or standards are more difficult to measure or evaluate. Again, they fall along the continue of being structural versus process oriented as mentioned in 9 above.

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

Chapter 5

Coordinated Program Monitoring, Differential Monitoring, Key Indicator, Risk Assessment, and Integrative Monitoring

This chapter demonstrates the national/federal initiatives addressing coordinated program monitoring. There are several excellent reports produced by the Administration for Children and Families (ACF), Office of Child Care (OCC), Office of Planning, Research and Evaluation (OPRE), Assistant Secretary's Office of Planning and Evaluation, and the Federal Department of Health and Human Services (HHS) which goes a long way in addressing this key issue. In any system where there are limited resources, we need to be as cost effective and efficient as possible. The handouts which accompany this text through the NARA Licensing Measurement course will provide you with many examples of how best to do this. These handouts/reports are all available of the RIKI and NARA websites as well as within the *NARA Licensing Curriculum (2000)*.

With a closed system and limited resources, a coordinated program monitoring system is critical to make certain that we have the necessary resources to effectively and efficiently protect the clients in the facilities we are mandated to license. The key term is "Do No Harm". The federal agency reports in this class will provide you with the parameters for building a program monitoring system that accomplishes this goal.

The lecture from the NARA licensing measurement course for this section of the chapter consists of a slide that builds upon *Caring for Our Children Basics (CFOCB*)(go to https://RIKInstitute.com) and how that publication came into

existence. Personally, I think it is one of the most significant publications (*CFOCB*) related to early care and education (ECE) standards development that has ever been produced.

CFOCB provides voluntary standards for all ECE to follow. It is the very essence of what coordinated program monitoring is all about in providing basic safeguards for all children while in out of home care.

Differential Monitoring, Risk Assessment, and Key Indicators

This ehandbook text has gotten into the details of differential monitoring, risk assessment, and the key indicator methodologies. We have tangentially addressed these methodologies throughout the text, but this chapter provides the step-by-step process of their development and implementation (see the following paragraphs). Also, there are several other publications that deal with this detail on the RIKI and NARA Websites.

The first step in utilizing the DMLMA (Differential Monitoring Logic Model and Algorithm) for a state is to take a close look at its Comprehensive Licensing Tool (CI) that it uses to collect violation data on all rules with all facilities in its respective state. If the state does not utilize a tool or checklist or does not review all violation data than it needs to consider these changes because the DMLMA is based upon an Instrument Based Program Monitoring System (IPM) which utilizes tools/checklists to collect data on all rules.

The second step for the state is to compare their state's rules with the National Health and Safety Performance Standards (*Caring for Our Children*)(available through https://RIKInstitute.com) to determine the overlap and coverage between the two. This is the first approach to validation which involves Standards review.

The third step for the state if it utilizes a Risk Assessment (RA) tool is to assess the relationship between this tool and Stepping Stones to determine the overlap and coverage between the two. This is a continuation of the first approach to validation which involves Standards review.

The fourth step for the state is to compare the results from the CI with the RA tools. This step is the second approach to validation which involves Measures. The correlation between CI and RA should be at the .50 level or higher (.50+).

In the fifth step, if a state is fortunate enough to have a QRIS – Quality Rating and Improvement System in place and has sufficient program quality (PQ) data available then they will have the ability to compare results from their CI tool with their PQ tool and validate outputs by determining the relationship between compliance with health and safety rules (CI) and program quality (PQ) measures, such as the ERS's, CLASS, CDPES, etc… This is a very important step because very few empirical demonstrations appear in the research literature regarding this relationship. This step is the third approach to validation which involves Outputs. It would be expected that lower correlations (.30+) would be found between CI and PQ because these tools are measuring different aspects of quality such as health & safety versus caregiver-child interactions or overall classroom quality.

The sixth step is for the state to generate a Key Indicator (KI) tool from the CI data base. Please see Fiene & Nixon (1985) and Fiene & Kroh (2000) publications available through the National Association for Regulatory Administration & Research Institute for Key Indicators Websites (https://RIKInstitute/com) for a detailed explanation of the methodology for generating a KI tool. This step is also part of the second approach to validation which involves Measures. The correlation between the CI and KI should be very high

(.70+) because the KI is a subset of predictor rules taken from the CI data base. If a state did not want to use the KI methodology, a direct comparison could be drawn from The *Thirteen Indicators of Quality Child Care* – this publication is available on the RIKI website.

The seventh step for the state is to use the RA and KI tools together to determine overall compliance of facilities and how often and which rules will be monitored for future visits. This is the basic component of a Differential Monitoring (DM) approach and continues the second approach to validation (Measures). Also, this step should drive decisions within the technical assistance/training/professional development (PD) system in what resources are allocated to a particular facility. It would be expected that moderate correlations (.50+) would be found amongst RA, KI, DM, and PD.

The eighth and final step for the state is to compare the results from the various monitoring tools (CI, PQ, RA, KI) with any child development outcome (CO) data they collect. This is a relatively new area and few, if any, states at this point have this capability on a large scale. However, as Early Learning Networks and Standards are developed, this will become more common place. This step is the fourth approach to validation which involves Outcomes. The correlations between CI, PQ, RA, KI and CO will be on the lower end (.30+) because there are so many other variables that impact children's development other than child care facilities.

Key Element Definitions: 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. 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.

Validation is a continuous approach and is not a once and done process. States should look at their monitoring systems on an on-going basis and make the necessary adjustments as data are collected and compared in order to keep program monitoring as cost effective and efficient.

In the readings/handouts, the learner will find several report examples which provide the details of the various methodologies. There are more than enough examples, so pick the ones you are most interested in seeing. For those of you who would like to see more, please go to the RIKI website and look under the report's webpage for additional examples.

Just as differential monitoring helped to change the landscape of program monitoring in making it more sensitive to targeted reviewing, integrative monitoring introduces a new conceptual lense in how program monitoring should be done. Just as coordinated monitoring focuses more on the type of care provided, integrative monitoring focuses on the actual standards and how they should be formatted. It is more of a delicate balancing act between regulatory compliance and quality programming when it comes to integrative monitoring which is very different from coordinated monitoring which emphasizes the facility type.

An interesting future research area is combining differential and integrative monitoring into a new approach to program monitoring which would emphasize risk assessment, key indicators, and quality programming into this new paradigm. Once this is done, it would be relatively easy to take those results and apply them within a coordinated monitoring approach. To a certain extent, ***Caring for Our Children Basics*** accomplishes this but without the increased emphasis on the quality programming side.

Integrative Monitoring

In the previous chapter, Chapter 4, the delineations between regulatory compliance and program quality were dichotomized showing how they were different and similar. This section of Chapter 5 introduces the notion of integrative monitoring where regulatory compliance and program quality are joined together in a cohesive management system.

The concept of integrative monitoring was introduced by Freer and Fiene (2023) in a ***Journal of Regulatory Science*** article in which they propose the concept along with specific constraints in implementing the approach.

In the past regulatory compliance and quality programming had their respective silos and were not integrated within the monitoring function. In fact, in ECE separate quality initiatives, such as Quality Rating and Improvement Systems (QRIS), accreditation, and professional development systems have been developed and implemented separate from licensing systems in most jurisdictions. In some cases, there were attempts to integrate the two arenas but generally these meant that there was acknowledgement that both existed but standards remained separated and assessors measuring compliance with the standards or rules were in different departments.

The Freer and Fiene (2023) model provides a unique means for combining regulatory compliance and quality programming into one comprehensive, effective, and efficient approach.

Freer & Fiene (2023). Regulatory compliance and quality programming: Constraints and opportunities for integration, ***Journal of Regulatory Science, Volume 11, Number 1***, pps 1-10.

Chapter 6

What Research Tells Us, What We Don't Know, and Examples

This ehandbook text has summarized what we know from the research literature about licensing measurement. There have been several advances in licensing measurement over the past couple of decades. Clearly the Regulatory Compliance Theory of Diminishing Returns has taken hold of policy development in licensing and regulatory administration. We have seen statutes change from requiring full 100% compliance in order to receive a license to operate to statutes that are requiring substantial regulatory compliance with all rules rather than full 100% compliance. Getting to those right rules rather than more or less rules. *Caring for Our Children: Basics* is an excellent example of this approach.

Another example is from developing countries, especially in Africa as so evidence by the work of researchers in Kenya who have utilized the theory of regulatory compliance repeatedly in various industries in order to develop and promulgate effective and efficient regulatory policy. In fact, we have a lot to learn from these initiatives because they are becoming part of a new way of doing policy research that is on a cutting edge.

Licensing key indicators and risk assessment rules are being used on a much larger scale as the differential monitoring/targeted monitoring approach has expanded. The latest Licensing Study conducted by NARA and the National Center for Early Childhood Program Quality has demonstrated that the majority of states are using one of these approaches.

The differential monitoring approach and its respective methodologies have gone through many enhancements in dealing with measurement and statistical nuances related to licensing data distributions, such as severe skewness, kurtosis, dichotomization of data groups, eliminating false negatives, limitations of nominal data analysis, moving from a nominal measurement scale to an ordinal measurement scale, identifying generic licensing key indicators, and the relationship between regulatory compliance & program quality (Chapters 3 & 4 highlighted this).

Eliminating false negatives had been a design issue in the key indicator methodology when substantial compliance is utilized for the high group dichotomization. It is not an issue when 100% full regulatory compliance is used but there will be instances when substantial compliance will have to be used. When this occurs a revision to the original methodology and algorithms has to be instituted which has been outlined in a RIKINotes Post (January 29, 2023). Utilizing this revision eliminates or at least mitigates the false negative effect.

All these above enhancements are basically dealt with and addressed in the *RIKI Technical Research Notes* found in the *ECPQIM/DMLMA text* as well as on the RIKI website Blog/Notes Page (https://RIKInstitute.com/blog/), or the National Association for Regulatory Administration website which was cited at the beginning of the eHandBook. The interested reader should find all these technical research notes in one of those venues. Just look towards the end of the webpage to find the research notes.

What Research Doesn't Tell Us

So, what are the gaps in the research related to licensing measurement that licensing researchers and regulatory scientists should be paying attention to? This text has provided some of the key gaps that have been identified to date. One area

for further research is the relationship between regulatory compliance and outcomes for clients. Are clients healthier and safer in highly compliant programs? Are we seeing fewer injuries in those programs of high regulatory compliance? This is a critical question that still needs definitive research and empirical evidence to confirm.

There still needs to be additional research that continues to validate the rules/standards selected, the measures themselves, and the relationship between regulatory compliance and QRIS systems. There has been considerable movement in the past decade with validation studies being completed in many states and provinces and this trend needs to continue. The results to date definitely appear to validate all these respective components in that they are working as expected, but I would feel more confident with additional replication studies being completed.

International, National, and State Examples

This section provides us with examples mainly through the specific tools that have been designed by different jurisdictions for the differential monitoring, key indicator and risk assessment methodologies described in this text. The readings and handouts provide many such examples which are available at https://RIKInstitute.com. You will find examples both from the USA as well as Canada. The methodologies have really taken off in the last decade as demonstrated by the number of contracts NARA has entered into with states and provinces throughout the United States (Montana, Michigan, Illinois, Indiana, Kansas, Florida, New York, Minnesota, California) and Canada (British Columbia, Alberta, Saskatchewan). Reports written describing these efforts are available on both the RIKI and NARA websites. See the graphic display at the end of this ebook in the Display, Graphs, Figures, and Charts Section.

All of these jurisdictions have demonstrated a certain consistency when it comes to licensing key indicator predictor rules and risk assessment rules. There are common themes that have emerged over the past 4 decades.

Here are key elements that should be present in a high-quality early care and education (ECE) program that any parent should be looking for when selecting their child care arrangement:

- Qualified ECE teachers.
- There is a stimulating and dynamic classroom environment where children are viewed as competent learners.
- A developmentally appropriate curriculum is used based upon the assessed individual needs of children.
- Opportunities for families and staff to get to know each other.
- Families receive information on their children's progress regularly using a formal process.
- Early childhood educators encourage children to communicate.
- Early childhood educators encourage children to develop reasoning skills.
- Early childhood educators listen attentively when children speak.
- Early childhood educators speak warmly to children

You will witness this consistency in the readings you have access to at https://RIKInstitute.com. Please check out the website because there are numerous publications and reports available to you. All the publications are in the public domain, so you are free to download them as you see fit.

The plan is to continue validating the methodologies to make certain that they are keeping children healthy and safe and are doing no harm. That is the key element of licensing measurement with a focus on health and safety similar to the

approach taken by the Nuclear Regulatory Commission (NRC) in keeping surrounding communities safe where nuclear power plants are located.

As has been repeatedly demonstrated in this ehandbook text, there is a delicate balance between regulatory compliance and program quality (remember chapter 4). Some industries are more geared towards the health and safety side of the equation while others seek a more balanced approach of regulatory compliance and program quality. I have attempted to address both in this text and hopefully have done an equally balanced approach in addressing both sides of the equation. It will be interesting to see how things play out as regulatory science continues to grow as a science and the impact of licensing measurement on the development of this very important science.

Chapter 7

Future Directions/Next Steps

This last chapter deals with where do we go from here. What are the next steps for licensing measurement. How do we combine the quantitative and the qualitative? How do we have a mixed methods approach? How do we combine the best aspects of regulatory compliance with program quality elements? Are there more effective ways to deal with terribly skewed data other than dichotomization? Does it make sense to move from a nominal to an ordinal measurement scale with regulatory compliance? All these are critical questions for the field of regulatory science and its accompanying licensing measurement. If we are truly going to build a science, we need to spend the requisite time on developing and implementing a solid scientific measurement strategy that is both reliable and valid.

Two of the most critical concepts that will need addressing are the ceiling effect/plateauing of quality data and the variance in the data distribution. The ceiling effect led to the regulatory compliance theory of diminishing returns and has had a major impact on the regulatory science field. Without a solution to this ceiling effect, it will continue to be difficult to distinguish between mediocre quality in programs and high quality in programs. At this writing, it appears that this ceiling effect is an inherent characteristic of regulatory compliance systems. This same ceiling effect does not appear to be present in quality-based program monitoring systems nor is the lack of variance present in those systems as well.

Lack of variance in data distributions can and should be addressed by building off weighting systems that look at risk or performance indicators. This can be a very effective and efficient approach to increase the variance in a regulatory

compliance data distribution. But it does need to be robust enough so that differences can be ascertained at the substantial regulatory compliance and the full regulatory compliance levels. It has always been easy to distinguish between low regulatory compliance and substantial regulatory compliance but that is not the case with substantial and full regulatory compliance when comparisons are made to program quality measures, such as the ERSs and CLASS tools.

And these two above issues lead us to another key balancing act between effectiveness and efficiency. As you have seen throughout this ehandbook text there are several concepts that need to be balanced with other domains in order to be both effective and efficient. A good example is the use of risk assessment and key indicators together when designing and implementing a differential monitoring system. However, it is always possible to put this delicate balance out of sync by placing too much emphasis on one or the other domain.

For example, a jurisdiction could become so efficient in utilizing key indicators that the tool has so few indicators that it begins to jeopardize the overall effectiveness of the monitoring system. And the other scenario is also a concern in which too many indicators are included on the tool in which effectiveness might increase but efficiency will decrease substantially. Finding that correct balance is an individual study into attempting various strategies where clients are not placed at additional risk but at the same time we don't want to jeopardize the overall quality of the program.

So this leads us to a balance between regulatory compliance and program quality which is the essence of the last sentence of the previous paragraph. It is also the essence of integrative monitoring which is attempting to focus both on regulatory compliance and quality. On the surface, this sounds really easy to develop a balance between regulatory compliance and program quality; but in reality, it is difficult to pull off. The

reason is that licensing and regulations are just not geared to deal with program quality. It is all about health and safety and focusing on risk aversion. So in moving forward this is going to be a difficult balancing act for most jurisdictions.

What are some of the other issues that we will need to address as we move forward? Validation studies are going to be key in moving forward as we determine if the monitoring systems we have designed are working as we intended. For the one size fits all, this probably will not be a heavy lift; but for differential monitoring and integrative monitoring I think these are going to be heavy lifts for most jurisdictions. These systems are not just descriptive based systems but have inferences built in and this is always more difficult to validate.

And in speaking of validation studies, a key validation study will be to validate the use of quality indicators which is the new kid on the block. Licensing key indicators have been around for some time, but quality key indicators have not and will need a good deal of research to determine what really are predictive indicators. I think we do have a good start based upon the studies that have been done with QRIS, accreditation, and professional development, our major quality initiatives in the USA. But additional research is still needed to validate the initial results. It is interesting to note that the first pilot testing and validation of quality key indicators has occurred in a Canadian Province and not in the USA. Our hats are off to the Province of Saskatchewan for being the first to pilot test and to validate this new approach.

And then there are the key measurement and statistical methods that need further development and refinement as it relates to licensing data. If we do move regulatory compliance attempting to balance it with program quality, there will most likely be experimentation in moving from a nominal measurement to an ordinal measurement scale. This idea has been suggested in the likes of a Regulatory Compliance Scale.

But it is still theoretical and has not been attempted yet. But for the future, this will become an important area of research.

On the statistical side, there will be the need to develop techniques to deal more effectively with skewed data distributions, false negatives, and other licensing data idiosyncracies. As I have said many times in this eHandBook licensing data are very unique. Part of this uniqueness is the fact that the data distributions are anything but normally distributed with very little variance. This is a major area of concern when it comes to statistical analysis and will need to be dealt with heads on in the coming years by regulatory scientists and licensing researchers.

This short ehandbook text is a first step in providing that scientific base for building a sound regulatory science, but I am hopeful that other licensing researchers and regulatory scientists build upon what has been presented and suggested in this eBook.

For those interested in pursuing any of these topics, please don't hesitate to go to the RIKI Institute or the NARA websites for additional detailed information. Here are the pertinent websites for your ease of access: https://rikinstitute.com
or https://www.naralicensing.org/key-indicators.

Research has been going on for approximately 50 years when the first kernels of what a regional model for monitoring would look like as it related to the human services, in particular early care and education. I never thought it would lead to its own statistical methodologies and altering how licensing and monitoring decision making would occur. And definitely did not think that "differential monitoring" would be referenced in Federal legislation with the re-authorization of CCDBG. And the regulatory science field which spans all industries and domains concerned with the application of rules and regulations

to our everyday existence has only coalesced over the last 20-25 years.

As I said earlier, the purpose of this ehandbook was as a short guide for those in the regulatory science and licensing research arenas to get a basic understanding of licensing measurement and program monitoring. By starting with it and using it in conjunction with all the publications and materials on the RIKI and NARA websites as well as the NARA Licensing Measurement course, it will provide an introduction to the state of the art regarding licensing measurement.

Let me leave you with a Regulatory Compliance Matrix (see table below) which summarizes the key points in this ehandbook when it comes to principles of regulatory compliance measurement, paradigms, and the quality continuum; but also points us in the direction for future research as these are the key elements for licensing measurement and monitoring systems.

The principles detail is provided in the appendix in the last document listed: *Ten Principles of Regulatory Compliance Measurement.* The paradigms detail can be found in chapter 2 of this ehandbook; while the quality continuum is in chapter 4. There is a good deal of overlap with the 10 principles, the 10 elements related to paradigms, and the 10 elements of the quality continuum. I would suggest focusing on these common elements and principles because they are the most significant pieces of the puzzle as it relates to regulatory compliance measurement. Taken together, these 30 principles and elements provide the basic parameters regulatory scientists, licensing researchers, and licensing administrators/policymakers should be focusing on when it comes to future research studies related to regulatory science.

Regulatory Compliance Matrix

Principles	***Paradigms***	***Quality Continuum***
Lack of variance	Substantial vs monolithic	Hard vs soft data
Ceiling effect	**One size fits all vs differential**	**Full vs partial compliance**
Difficulty between full and high	Rules are equal vs not equal	Rules vs indicators
Nominal measurement	**Do things well vs do no harm**	**Do no harm vs do good**
Moving nominal to ordinal	**Strength based vs deficit**	Open vs closed system
Dichotomization	Formative vs summative	**Structural vs process quality**
Lack of reliability and validity	**Program quality vs compliance**	**Risk vs performance**
Skewed data	**100-0 vs 100 or 0**	**Nominal vs ordinal**
Ease between high and low	QRIS vs licensing	Gatekeeper vs enabler
False negatives	Linear vs non-linear	**Ceiling effect**

Appendices

- **The Relationship between Early Care & Education Quality Initiatives and Regulatory Compliance**
- **Regulatory Compliance, Licensing, and Monitoring Measurement Principles: Rule Compliance Versus Rule Performance**
- **What is the Relationship between Regulatory Compliance and Complaints in a Human Services Licensing System?**
- **The Implications in Regulatory Compliance Measurement When Moving from Nominal to Ordinal Scaling**
- **So Which Is Better: Differential Monitoring & Abbreviated Inspections or Comprehensive Inspections?**
- **The Dichotomization and Bi-Polarization of the Matrix Data Base**
- **Enhanced Dichotomization Model for Generating Licensing Key Indicator**
- **The Relationship of Licensing, Head Start, Pre-K, QRIS, Accreditation, and Professional Development and their Potential Impact on Child Outcomes**
- **Policy Commentary: Regulatory Science Measurement Issues of Skewness, Dichotomization of Data, and Nominal versus Ordinal Data Measurement**
- **A Potential Reason for Skewed Regulatory Compliance Data Distribution**
- **Data Distribution in Regulatory Science**
- **Ten Principles of Regulatory Compliance Measurement**

These above appendices provide additional detail to the above chapters in delving deeper into some of the key points made in Chapters 1-7.

References follow the appendices listing pertinent publications related to licensing measurement and monitoring systems.

Figure, Charts, Graphs, and Displays supporting the text in this eHandBook follow the Reference Section.

The Relationship between Early Care & Education Quality Initiatives and Regulatory Compliance

Over the past couple of decades there has been many early care and education initiatives, such as Quality Rating and Improvement Systems (QRIS), Professional Development, Training, Technical Assistance, Accreditation, and Pre-K programs to just name a few. Validation and evaluation studies have begun to appear in the research literature, but in these studies, there has been few empirical demonstrations of the relationship between these various quality initiatives and their impact on regulatory compliance or a comparison to their respective regulatory compliance. This brief technical research note will provide examples of these comparisons taken from the Early Childhood Program Quality Improvement and Indicator Model (ECPQI2M) Data Base maintained at the Research Institute for Key Indicators (RIKIllc).

I have written about this back in 2014 (Fiene, 2014) in how the various quality initiatives were having a positive impact on the early care and education delivery system but at that point regulatory compliance data were not available. Today, in 2019, with many changes and developments in state data systems, this is no longer the case. Now it is possible to explore the relationships between data from the various quality initiatives and licensing. Several states in multiple service delivery systems have provided replicable findings in which I feel comfortable reporting out about the relationships across the data systems.

What we now know is that there is a positive and statistically significant relationship between regulatory compliance and moving up the QRIS Quality Levels. In other words, facilities have higher compliance in the higher QRIS Quality Levels and lower compliance in the lower QRIS Levels or if they do not

participate in their state's respective QRIS ($F = 5.047 – 8.694$; $p < .0001$). Other quality initiatives, such as being accredited, shows higher compliance with licensing rules than those facilities that are not accredited ($t = 2.799$ - 3.853; $p < .005$ - $.0001$).

This is a very important result clearly demonstrating the positive relationship between regulatory compliance and quality initiatives. I have some additional state data sets that I will add to the ECPQI2M data base and will continue to analyze these relationships.

Regulatory Compliance, Licensing, and Monitoring Measurement Principles: Rule Compliance Versus Rule Performance

The purpose of this short paper is to delineate the parameters of regulatory compliance, licensing and monitoring measurement principles (throughout this paper the term "regulatory compliance" will be used to encompass these principles). Regulatory compliance is very unique when it comes to measuring it because it is very different from other measurement systems and this impacts how one uses various statistical analyses. In this paper, the limitations of the measurement system will be highlighted with potential solutions that have been devised over the past several decades. Hopefully this paper will add to the measurement and statistical analysis licensing research literature. It is meant for those agency staff who are responsible for designing regulatory compliance, licensing and monitoring systems. Its focus is the human services but the basic principles can be applied to any standards-based system that is based upon a compliance or performance model.

The organization of this paper is as follows. First, let's introduce what is included when we talk about measurement principles for regulatory compliance, licensing and monitoring systems. Second, provide examples that should be familiar to most individuals who have been involved in the human services, in particular the early care and education field. Third, what are the limitations of these various systems that have been identified in the research literature. Fourth, what are some potential solutions to these limitations. And, fifth, what are the next steps and where do we go to build reliable and valid measurement systems dealing with regulatory compliance, licensing, and program monitoring as these relate to the human services delivery system.

So, what is included in this approach. I can be any rule, regulation, or standard based measurement system. Generally, these systems are focused on a nominally based system, sometimes they will be ordinally based. By a nominally based system, either the facility being assessed is in compliance with a particular set of rules, regulations, or standards or it is not. In an ordinally based system, a facility may attain a score on a Likert scale, such as 1 through 5 where 1 is non-optimal and 5 is excellent. These types of measurement scales involve a performance component and are not limited to more of a compliance focus as is the case with a nominally based system. These distinctions are important as one will see later in this paper when it comes to the selection of the appropriate statistics to measure data distributions and the subsequent analyses that can be undertaken.

What are examples of these types of systems? For nominally based systems, just about all the licensing systems in the USA, Canada and beyond employ this type of measurement strategy. As has been said in the previous paragraph, either there is compliance or there is not. It is very black or white, there are not shades of gray. For ordinally based systems, these systems are a bit more diverse. Accreditation, Quality Rating and Improvement Systems (QRIS), the new Head Start Grantee Performance Management System (GPMS), the Environmental Rating Scales, and the CLASS are all examples of ordinally based systems based upon a Likert type measurement system. There are many others, but as a research psychologist whose total career (50 years) has been spent in early care and education, this has been the focus of my research.

The limitations of the above systems are numerous and, in some ways, are difficult to find solutions. In the past, these measurement systems have focused more on the descriptive aspects of data distributions rather than attempting to be predictive or inferential. The first major limitation of the data from regulatory compliance systems is the fact that the data distribution is markedly skewed. What does skew data mean?

Most data distributions are normally distributed with very few occurrences at the extremes with the majority of the cases in the middle section of the measurement scale. IQ is an example of a normally distributed data distribution. In a skew data distribution, the majority of data are at one end of the data distribution, either at the positive end or the negative end of the distribution. With regulatory compliance data, it is at the positive end with the majority of facilities being in full or 100% compliance with the rules. Very few of the facilities are at the negative end of the distribution.

What is the big deal? The big deal is that statistically we are limited in what we can do with the data analyses because the data are not normally distributed which is an assumption when selecting certain statistical tests. Basically, we need to employ non-parametric statistical analyses to deal with the data. The other real limitation is in the data distribution itself. It is very difficult to distinguish between high and mediocre facilities. It is very easy to distinguish between high and low performing facilities because of the variance between the high performing facilities and the low performing facilities. However, that is not the case between high and mediocre preforming facilities. Since the majority of facilities are either in full or substantial compliance with the rules, they are all co-mingled in a very tight band with little data variance. This makes it very difficult to distinguish differences in the facilities. And this only occurs with regulatory compliance data distributions. As will be pointed later in this paper, this is not the case with the second measurement system to be addressed dealing with ordinal measurement systems.

There is also a confounding factor in the regulatory compliance data distributions which has been termed the theory of regulatory compliance or the law of regulatory compliance diminishing returns. In this theory/law, when regulatory compliance data are compared to program quality data, a non-linear relationship occurs where either the facilities scoring at the substantial compliance level score better than the fully

compliant facilities or there is a plateau effect and there is no significant difference between the two groups: substantial or fully compliant facilities when they are measured on a program quality scale. From a public policy stand point, this result really complicates how best to promulgate compliance with rules. This result has been found repeatedly in early care and education programs as well as in other human service delivery systems. It is conjectured that the same result will be found in any regulatory compliance system.

Another limitation of regulatory compliance data is the fact that it is measured at a nominal level. There is no interval scale of measurement and usually not even an ordinal level of measurement. As mentioned above, either a facility is in compliance or not. From a statistical analytical view, again this limits what can be done with the data. In fact, it is probably one of the barriers for researchers who would like to conduct analyses on these data but are concerned about the robustness of the data and their resulting distributions.

Let's turn our attention to potential solutions to the above limitations in dealing with regulatory compliance data. One potential solution and this is based upon the theory of regulatory compliance in which substantial compliance is the threshold for a facility to be issued a license or certificate of compliance. When this public policy determination is allowed, it opens up a couple of alternate strategies for program monitoring and licensing reviews. Because of the theory of regulatory compliance/law of regulatory compliance diminishing returns, abbreviated or targeted monitoring reviews are possible, differential monitoring or inferential monitoring as it has been documented in the literature. This research literature on differential monitoring has been dominated by two approaches: licensing key indicators and weighted risk assessments.

A second solution to the above limitations deals with how we handle the data distribution. Generally, it is not suggested to dichotomize data distributions. However, when the data

distribution is significantly skewed as it is with regulatory compliance, it is an appropriate adjustment to the data. By essentially having two groups, those facilities that are in full compliance and those facilities that are not in full compliance with the rules. In some cases, the fully compliant group can be combined with those facilities that are in substantial compliance but this should only be employed when there are not sufficient fully compliant facilities which is hardly never the case since population data and not sampled data are available from most jurisdictions. When data samples were drawn and the total number of facilities were much smaller, substantial compliant facilities were used as part of the grouping strategy. The problem in including them was that it increased the false negative results. With them not being included, it is possible to decrease and eliminate false negatives. An additional methodological twist is also to eliminate and not use the substantial compliant facilities at all in the subsequent analyses which again helps to accentuate the difference scores between the two groups of highly compliant and low compliant scoring facilities.

The next steps for building valid and reliable regulatory compliance systems are drawing upon what has been learned from more ordinally based measurement systems and applying this measurement structure to regulatory compliance systems. As such, the move would be away from a strict nominally based measurement to more ordinal in which more of a program quality element is built into each rule. By utilizing this paradigm shift, additional variance should be built into the measurement structure. So rather than having a Yes/No result, there would be a gradual Likert type (1-5) scale built in to measure “rule performance” rather than “rule compliance” where a “1” indicates non-compliance or a violation of the specific rule. A “5” would indicate excellent performance as it relates to the specific rule. A “3” would indicate compliance with the specific rule meeting the specifics of the rule but not exceeding it in any way.

This paradigm shift has led to the creation of Quality Rating and Improvement Systems (QRIS) throughout the USA because of a frustration to move licensing systems to more quality focused. The suggestion being made here is to make this movement based upon the very recent developments in designing such systems as is the case with Head Start monitoring. Head Start GPMS is developing an innovative Likert based ordinal system which incorporates compliance and performance into their monitoring system. Other jurisdictions can learn from this development. It is not being suggested as a replacement for QRIS or accreditation or ERS/CLASS assessments but as a more seamless transition from licensing to these various assessments. As indicated by the theory of regulatory compliance and the law of regulatory compliance diminishing returns, this relationship between licensing and program quality is not linear. By having this monitoring system approach in place, it may be able to reintroduce more of a linear relationship between licensing and program quality.

What is the Relationship between Regulatory Compliance and Complaints in a Human Services Licensing System?

Within licensing measurement and the validation of licensing systems it is particularly difficult to have specific outcome metrics that can be measured within a human services licensing system. The purpose of this technical research note is to propose a potential solution to this problem.

Probably the most accurate measures of licensing outcomes focuses on improvements in the health and safety of clients within human services licensed facilities, such as: fewer injuries (safety) or higher levels of immunizations (health). Another measure related to client satisfaction is the number of complaints reported about a licensed facility by clients and the general public. The advantage of using complaints is that this form of monitoring is generally always part of an overall licensing system. In other words, the state/provincial licensing agency is already collecting these data. It is just a matter of utilizing these data in comparing the number of complaints to overall regulatory compliance.

The author had the opportunity to have access to these data, complaint and regulatory compliance data in a mid-Western state which will be reported within this technical research note. There are few empirical demonstrations of this relationship within the licensing research literature. The following results are based upon a very large sample of family child care homes (N = 2000+) over a full year of licensing reviews.

The results of comparing the number of complaints and the respective regulatory compliance levels proved to show a rather significant relationship ($r = .47$; $p < .0001$). This result is the first step in attempting to understand this relationship as well as

developing a methodology and analysis schema since directionality (e.g., did the complaint occur before or after the regulatory compliance data collection?) can play a key role in the relationship (this will be developed more fully in a future technical research note). The focus of this research note was to determine if any relationship existed between regulatory compliance and complaint data and if it is worth pursuing.

It appears that looking more closely at the relationship between complaint and regulatory compliance data is warranted. It may provide another means of validating the fourth level of validation studies as proposed by Zellman and Fiene's OPRE Research Brief (Zellman, G. L. & 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 2012-29. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services) in which four approaches to validation are delineated for Quality Rating and Improvement Systems (QRIS). This author has taken this framework and applied it to licensing systems (Fiene (2014). Validation of Georgia's Core Rule Monitoring System, Georgia Department of Early Care and Learning) and more recently proposed as the framework for Washington State's Research Agenda (Stevens & Fiene (2018). Validation of the Washington State's Licensing and Monitoring System, Washington Department of Children, Youth, and Families).

The Implications in Regulatory Compliance Measurement When Moving from Nominal to Ordinal Scaling

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

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

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

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

By employing this alternate paradigm, it is a first demonstration of the use of the Key Indicator Methodology in both licensing and quality domains. The Key Indicator Methodology has been utilized a great deal in licensing but in few instances in the program quality domain. For example, over the past five years,

I have worked with approximately 10 states in designing Licensing Key Indicators but only one state with Quality Key Indicators from their QRIS – Quality Rating and Improvement System. This new paradigm would combine the use in both. It also takes advantage of the full ECPQI2M – Early Childhood Program Quality Improvement and Indicator Model by blending regulatory compliance with program quality standards.

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

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

So Which Is Better: Differential Monitoring & Abbreviated Inspections or Comprehensive Inspections?

During 2019 and 2020, several validation studies have been or are being completed in the states of Washington, Indiana, and in the Province of Saskatchewan. These validation studies are determining if the key indicator and risk assessment methodologies are valid approaches to conducting abbreviated inspections in comparison to more comprehensive inspections in which all rules are assessed. These abbreviated inspections are a form of differential or targeted monitoring. This technical research note focuses on the empirical evidence to determine the efficacy of these approaches, are they better than doing comprehensive reviews when it comes to health and safety outcomes.

When the key indicator and risk assessment methods were originally proposed in the 1980's, an outcome validation study was completed in Pennsylvania during 1985 – 1987 by Kontos and Fiene to determine what impact those methods had on children's development. In that original study, it was determined that the Child Development Program Evaluation Indicator Checklist (CDPEIC) was more effective and efficient in predicting child development outcomes than the more comprehensive Child Development Program Evaluation. In fact, the CDPEIC and the accompanying Caregiver Observation Scale (COFAS) were as effective and more efficient than the ECERS – Early Childhood Environmental Rating Scale in that study.

Fast forward to 2019 – 2020, in the province of Saskatchewan, Canada, and a similar study was undertaken but in this case the outcomes were more based upon health and safety rather than child development developmental outcomes. In this case, again the key indicator and risk assessment tool was both a more effective and efficient model over the more comprehensive

inspection approach giving credence to utilizing differential monitoring with abbreviated inspections.

In both of the above validation studies involving either child development assessment outcomes or health & safety outcomes, a 16 to 28% increase in effectiveness was observed in the outcome data. In the abbreviated or targeted inspections, 33% of the total rules or less are used to make the determination of regulatory compliance. It is like having the best of both worlds when it comes to effectiveness (16 – 28% increase in outcomes) and in efficiency (66% fewer rules being used). These studies help to validate the use of differential monitoring as a viable alternative to the more comprehensive one-size-fits-all monitoring reviews.

The Dichotomization and Bi-Polarization of the Matrix Data Base

This latest technical note updates the thresholds for the high and low groups within the key indicator matrix. This technical note is based upon the latest studies during the early 2015 time frame in which very large data distributions were available to test certain criteria with the key indicator methodology. Because of the extreme skewness present in licensing/regulatory data, certain statistical adjustments need to be made so that the analyses performed reflect the distribution of data. One of these statistical adjustments is the dichotomization of data which is generally not suggested with the exception of very skewed data. Since licensing data are so skewed, this adjustment has been used throughout the key indicator methodology. However, an additional adjustment is now warranted given not only the skewness of data but also because of the data being nominal in nature. This additional adjustment I am calling the bipolarization of data in order to accentuate the differences between the high and low groups within the key indicator matrix.

I have tested several data sets utilizing bi-polarization and found that the results are more significant with its use than without its use. Please keep in mind that licensing data is very different from other forms of data found in the early care and education (ECE) research literature. It is not like the ERS or CLASS data which is more normally distributed and lends itself to more parametric statistical analyses. Licensing data are nominal in nature and always very skewed which means that more non-parametric methods are warranted, such as phi coefficient and dichotomization of data. An example of how this actually works may help.

Licensing data are measured as either being in or out of compliance. There is no middle ground, it is not measured on a Likert scale. Therefore it is nominal in nature, either it is all

there or it is not. Licensing data are also measured in the sense that all rules are created equally, in other words, they all have the same weight or importance, such as 1 = compliance; 0 = non-compliance. Being in full 100% compliance which means 0 violations is the goal of a regulatory/licensing system. One does not want to see many violations of the rules because this will place children at risk of harm and the purpose of an early care and education (ECE) licensing/regulatory system is to reduce the potential harm to children. In the licensing measurement literature, this 100% compliant group is generally labeled or considered the high compliant group. With some licensing laws which allow substantial but not full 100% compliance with the full set of rules, it would then be allowable to have possibly 1 or 2 violations and still be considered in this high compliant group. The low compliant group has been generally any program that had any non-compliance or had 2 or more violations. When these two groups were compared to each individual rule utilizing the phi coefficient formula it was found that a more accurate approach was to accentuate or increase the difference between the high and low groups by eliminating the intervening violations in following manner: high group of 0 violations; 1-4 violations being eliminated; 5+ violations defined as the low group. This additional bi-polarization of data helped to accentuate the differences in calculating the phi coefficient and provided a more sensitive key indicator tool.

Another data distribution issue that should be addressed here that justifies the above cutoffs is that there is very little variance in licensing/regulatory data. Generally the frequency distribution is 20 or less and the average set of rules is over 200 rules. So the frequency distribution is extremely skewed within less than 10% of the potential data distribution. Also, the majority of programs are 100% in compliance with all the rules. And an additional complication is that the scoring of each rule is scored as if it had an equal risk value when in reality the rules can place children at either great risk to relatively little risk if found non-compliant. These measurement issues are very

different than in other measurement systems such as ERS or CLASS. The important message to take from this is that rules are not a ruler, they do not measure things equally and cannot be analyzed or compared to other measurement systems that are more normally distributed.

Although licensing is part of the program quality continuum in establishing basic health and safety standards for children, it is a system with measurement limitations that can only be compared on a nominal basis making several statistical adjustments as suggested above necessary.

Enhanced Dichotomization Model for Generating Licensing Key Indicators

The licensing key indicator methodology has been evolving over the past decade in making it more sensitive to the selection process of the specific rules to be included as key indicators. Some of the enhancements can occur because of state licensing data systems being able to provide population data rather than having to select sample data. Because of the nominal nature of licensing data and the severe skewness of the data distributions, non-parametric statistical approaches need to be employed in the analysis of the data.

A key component in the analysis of the licensing data distributions is to dichotomization of the data which is generally not warranted but is acceptable with very skewed data distributions. The dichotomization that has been most successful is a H25/M50/L25 distribution in which H25 represents the High Group of regulatory compliance, M50 which represents the Mediocre or Middle Group of regulatory compliance, L25 which represents the Lowest Group of regulatory compliance. In the past, the methodology allowed for full and substantial compliance within the High Group. This decision is no longer recommended. Rather, in order to decrease the number of False Negatives, it is now recommended that only Full (100%) regulatory compliance is used in defining the High Group. This eliminates the possibility of False Negatives.

By making this above change and in using the full distribution of licensing data, it enhances the results for generating the licensing key indicator rules. For additional information on this modeling please see: Fiene, Richard (2018), "***ECPQIM National Data Base", Mendeley Data, V1. http://dx.doi.org/10.17632/kzk6xssx4d***.1 This data base provides the detailed ECPQIM data distributions for the above changes. The enhancements increase the phi coefficients and reliability in either moving or not moving from abbreviated

inspections to full comprehensive inspections. This data base also contains clear demonstrations of the efficacy of the ECPQIM – Early Childhood Program Quality Improvement and Indicator Model as a vehicle for improving early care and education programs.

The Relationship of Licensing, Head Start, Pre-K, QRIS, Accreditation, and Professional Development and their Potential Impact on Child Outcomes

This short paper will provide some thoughts about the various public policy initiatives/systems to improve early care and education, such as licensing, Head Start, Pre-K, QRIS, accreditation, and professional development and their potential impact on child outcomes. Early care and education is at a major crossroads as a profession in attempting to determine which quality initiatives have the greatest impact on children. Results are starting to come in from early studies which may provide some guidance as policy makers begin making decisions about where to focus their limited funding resources.

Improving early care and education programs has a long public policy history as we attempt to find the most cost effective and efficient means for attaining this lofty goal. There have been many ups and downs over the years where funding was adequate and when it was not, but our desire to accomplish this goal has always been front and center. Now, as a profession, we are at somewhat of a cross-roads in determining which of the many quality initiatives appear to have the greatest impact on children's development. When I refer to children's development, I am looking at the whole child from the perspective of a child's developmental status as well as the child's health and safety.

Presently we have many quality initiatives to look at which is a very good thing since at times in the past we did not always have so many choices. Probably the one constant throughout the history of early care and education in the past century has been licensing or regulations/rule formulation. Some many argue that licensing is not a quality initiative but I would suggest that licensing has many of the structural aspects of quality that have been identified in the research literature. The other quality initiatives I will discuss have really started and been

implemented in the very later part of the 20th century so we are talking about a relatively new science when we think about having its intended impact on children. Also, I am talking about large public policy initiatives rather than highly structured, single focused research studies involving small samples of children.

Let's start with licensing since this system has been present for the longest period of time. The purpose of licensing is to act as the gatekeeper to the early care and education field in which only those providers who meet specific standards, generally called rules or regulations are permitted to operate and care for children. The rules are dominated by health and safety concerns with less emphasis on curriculum planning and staff-child interactions. The rules measure more structural aspects of quality than the process aspects of quality; dealing with what attorney's call the "hard data" rather than the "soft data".

Since licensing rules allow entry into the early care and education field to provide services usually the rules are not overally stringent with the majority of providers being in high compliance if not full compliance with all the rules. This would be expected since these are basic health and safety standards. And in fact when one looks at compliance data, it is extremely skewed with the majority of providers having very high compliance scores with relatively few violations of the rules. However, this does introduce a certain difficulty in using these data for decision making purposes at an aggregate level because so many providers score at a high level it becomes increasingly difficult to distinguish between the really excellent providers and the somewhat mediocre providers. Another way of looking at this skewing of the data is to term it as a plateau effect in which there is very little variance at the upper ends of the compliance spectrum. This is a major issue with skewed data and basic standards which is an important consideration with licensing but will also be an important consideration when one looks at the other quality initiatives to be addressed shortly.

Because of this plateau effect with licensing data, it may explain much of the lack of relationships found between compliance with rules and any types of outcomes related to children's outcomes and provider's overall quality. However, with licensing data and making comparisons to children's outcomes we should be looking at general health data such as immunization status and safety data such as the number of injuries at programs with varying levels of compliance with health and safety rules.

A significant development over the past two decades has been the development of national health and safety standards with the publication of Caring for Our Children (CFOC3) and Stepping Stones (SS3). Although these standards are not required but are only recommended practice that provides guidance to states as they revise their rules, these two documents have been embraced by the licensing/regulatory administration field. Although unlikely, if not impossible, to comply with all the CFOC3 standards, it would be interesting to compare states on this set of standards which may add a good deal of variance to the basic health and safety data that has been missing with licensing rules.

The next system to look at is the national Head Start program. Out of the major programs that are national in scope, Head Start has a long history of providing services to low-income children and their families. Head Start Performance Standards are definitely more stringent than licensing rules but not as stringent as accreditation standards. Based upon Head Start's more stringent standards and the additional supports that are part of its program, Head Start generally scores higher on program quality tools (e.g., CLASS or ERS) than licensed child care in states.

With Head Start programs, we at times find skewing or plateauing of data when we compare compliance with the Head Start Performance Standards (HSPS) and program quality tools such as the CLASS. However, this is dependent upon the

various subscales within the CLASS in which the plateauing of data does not occur all of the time. I think that has a lot to do with the HSPS being fairly stringent standards as compared to state licensing rules in general.

A program that has gotten a good deal of support at the state level are Pre-K programs. These programs come with stricter standards than licensed child care with an emphasis on the professional development of staff. There is more concern about the process aspects of quality which focus more on teacher-child interactions. This emphasis on teacher-child interaction has paid off in which these programs generally are high performers when you compare Pre-K funded classrooms to licensed child care classrooms. In fact, Pre-K funding appears to have a positive impact on licensed child care in raising overall quality scores on the ECERS-R for all classrooms in programs that receive Pre-K funding even if some of the classrooms are not the direct beneficiaries of the funding. This is a very significant finding because we knew that Pre-K funding increased the quality of care in classrooms receiving those funds, but now, it appears that there is a spillover effect to all classrooms co-located with Pre-K funded classrooms. I must admit that I was initially skeptical when Pre-K funding was first proposed because I thought it would take funding and the focus away from improving licensed child care at the state level; but it appears that the advocates for Pre-K were right in their assertion that Pre-K would increase the quality of all early care and education which includes licensed child care.

A more recent entry into the state funding scene are QRIS (Quality Rating and Improvement Systems) which build upon licensing systems, are voluntary, and have substantial financial incentives for participating in this quality improvement system. It is too early to really determine if QRIS is having the intended impact because the program is so new (50% of states have a QRIS), and the penetration rate is usually below 50% in any given state (remember the system is voluntary). However, in the few studies done, the results are mixed. It does appear that

programs which move up the various star levels do increase the quality of care they provide; but in a most recent study looking at child outcomes, no relationship was found between increasing levels of compliance with QRIS standards and how well children did in those programs with the exception of CLASS scores in which teacher-child interactions were measured and emphasized – here there were significant relationships between higher scores on the CLASS and child outcomes.

Accreditation systems come in many varieties but there are only three that I know of in which empirical studies have been done to validate their systems: NAEYC, NECPA for centers and NAFDC for homes. Also reliability testing has been done in each of these systems. Accreditation is a rigorous self-study that really improves programs through the self-study process. This should come as no surprise because we have known for some time that program monitoring all by itself leads to program improvements. Now when you couple that with technical assistance you see even more improvement. Accreditation is usually the other pillar of a QRIS system with licensing being the first pillar. The QRIS standards fill the gap from licensing to accreditation. Accreditation is a voluntary system just as in most cases with QRIS. However, in accreditation we are reaching less than 10% of the programs with the majority of these attaining NAEYC accreditation. NECPA and NAFDC have much smaller market shares.

The last system to be addressed is the professional development systems that have been established in all states. This is one quality improvement initiative that has 100% penetration in all states. It is usually tied to QRIS through technical assistance and mentoring (coaching). When it focuses on mentoring rather than workshops, it has demonstrated its effectiveness in changing teachers behaviors in how they interact with children in their care in a very positive fashion. This is very important because the research literature is clear about the importance of the teacher-child interaction when it comes to child outcomes.

Professional development runs the gamut from pre-service (University based programs) to in-service (training, technical assistance, mentoring, coaching) programming for teachers and directors.

So where does this leave us when policy makers begin to try to determine which quality improvement initiatives should be invested in to start with, which to increase in funding, and maybe even which ones should be defunded. I think there are some trends we need to begin to look at, such as the following:

1) Having stringent and rigorous standards is very important. The more that we do not, the more opportunities for mediocre programs to score artificially higher on whatever scale that is used. This is evident with licensing data where the data are significantly skewed with a major plateau effect at the upper end of compliance rules/regulations.

2) Emphasis on teacher-child interaction needs to be paramount in our quality improvement initiatives. Working with teachers through mentoring/coaching appears to be most effective in changing teachers' behaviors in interacting more positively with children.

3) Making sure we are measuring the right outcomes. Match health and safety standards with health and safety outcomes for children. Match developmental outcomes for children with standards that emphasize positive teacher-child interactions.

4) Building upon #1 above, find what the key indicators are with all the data that we collect. We are spending too much time in looking at too many things which in many cases are simply just not the right things to look at. As states' data systems become more sophisticated, and they are, this will be easier to do. Let's begin to utilize the data we have already collected.

Policy Commentary: Regulatory Science Measurement Issues of Skewness, Dichotomization of Data, and Nominal versus Ordinal Data Measurement

The purpose of this policy commentary is to provide some context for regulatory scientists in pursuing regulatory policy analysis, especially as it relates to regulatory compliance and human service licensing data. Regulatory scientists have dealt with non-parametric data in the past but in dealing with regulatory compliance and human service licensing data are just so different from previously measured data in that the nature of the data is nominal and extremely skewed to the point that several adjustments need to be made in order to analyze the data.

Although the examples being referred to in this policy commentary are from the human services field and discipline, I am certain that many of the basic concepts presented will pertain to other disciplines and fields of study that are impacted by regulatory science. These concepts are not unique to a particular discipline but rather are unique to regulatory science which has particular parameters, concepts, and truths which are pertinent to how regulations/rules/standards are formulated and then implemented in various jurisdictions or disciplines.

There are very logical reasons why regulatory compliance and licensing data are so extremely skewed. These data represent compliance with basic health and safety rules and regulations which provide the basic safeguards for children, youth, and adults while being cared for in a form of human services, such as child care, youth residential, or adult assisted living care.

Very honestly a state agency would not want to find their regulatory compliance data being normally distributed because this would be an indication that the facilities were in low compliance with the state's rules and regulations. Having the regulatory compliance data be highly negatively skewed is

actually a good result from a public policy standpoint but not from a statistical analytical standpoint. Having 50-60% of your scores within a three-to-five-point range when there may be as many as 300-400 data points leaves very little variance in the data. It also leads to being very difficult to distinguish between the high performers and the mediocre performers. This finding has led to a theory of regulatory compliance in which substantial compliance but not full compliance with all rules and regulations is in the best interests of the clients being served (Fiene, 2019).

In the regulatory science field, this has led to public policies emphasizing substantial compliance in order to be a licensed human service facility, such as a child care center, youth residential program, or an adult assisted living center. The other aspect of regulatory compliance and licensing data for regulatory scientists to consider is that the data are nominal in measurement, either a facility is in compliance or out of compliance with a specific rule or regulation. There are no gray areas, no measurement on an ordinal scale.

There has been some discussion in the regulatory science field for the use of weighted risk assessment methodologies which could introduce more variance in the data based upon the assumption that all rules or regulations are not created equal nor are they administered equally (Stevens & Fiene, 2019). Another discussion revolves around the introduction of more program quality into the basic health and safety rules and regulations that could extend the nominal compliance determination to an ordinal scale that goes beyond the basic compliance level (Fiene, 2018).

These measurement idiosyncrasies of regulatory compliance and licensing data are presented for regulatory scientists to consider if they begin to analyze public policies that involve basic health and safety rules and regulations which are very different from other public policies being promulgated by state and national governments. For the interested reader, an

international data base for regulatory compliance and human services licensing data has been established and maintained by the Research Institute for Key Indicators and Penn State University over the past 40 years at the following URL - (http://RIKInstitute.com)

However, the hope is that other disciplines will begin to look at their data more closely to determine the natural data distributions and ascertain if they are equally as skewed as has been found in human service regulatory data. Are you measuring the data at a nominal level? Could they be measured at an ordinal level based upon a Likert scale? The data being referred to are regulatory compliance data which are pegged to specific rules/regulations/standards. It is not based upon other types of data collected within a regulatory frame of reference, such as basic demographic or descriptive data.

A Potential Reason for Skewed Regulatory Compliance Data Distributions

One thing that is ever present with regulatory compliance data distributions is that they are terribly skewed. See the previous post which provides a definition of skewed distributions and their implications. This post is going to attempt to provide a potential answer to why the data base is skewed.

At first, I was led to believe that potentially the skewness in the data was a result of the rules not being stringent enough, in other words, the health and safety standards were too easy to comply with. That could definitely be a contributing factor but this is not the case in all instances when one compares state human service rules and regulations and the Head Start Performance Standards. I think a much deeper structure may be operating that is more philosophical rather than practical.

The philosophy of regulatory compliance and rule formulation is one of risk aversion. In other words, how do we mitigate risk that potentially increases the chances of mortality or morbidity in the clients being served when a specific rule is out of compliance. This philosophy emphasizes the elimination of a risk, taking something away rather than adding to it. It is essentially, "Do No Harm". It is interesting to note that generally regulatory compliance scoring is nominal in being either "Yes" or "No"; and a lower score is better than a higher score, there are fewer violations of rules. Not the way most assessment tools are designed.

For example, when one looks at program quality, this system is based upon the open-endedness adding to rather than taking away. It is all about, "Do Good" rather than "Do No Harm". Generally when you look at the data distributions here, they are more normally distributed without the skewed nature of regulatory compliance data distributions. Generally program quality scoring is ordinal in nature on a Likert Scale. A higher

score is better than a lower score. Makes sense in that when you have more of a good thing, the higher the score. And the philosophy of program quality is one of improvement with relatively little emphasis on risk aversion.

This is an alternate explanation to why regulatory compliance data distributions are so terribly skewed in comparison to other program quality measures.

Data Distributions in Regulatory Science

Data distributions in the human services as they relate to regulatory compliance are generally very skewed distributions which means that the majority of facilities being assessed/inspected will usually fall very close to the 100% compliance level. There will also be an equally large number of facilities that are in substantial regulatory compliance (99% - 98% compliance levels). And then there are much fewer facilities that are either at a mid or low level of regulatory compliance (97% or lower compliance levels). One might say that getting a score of 97% on anything doesn't sound like it is mediocre or low but keep in mind we are addressing basic health and safety rules and not quality standards. So having several health and safety rules out of compliance is a big deal when it comes to risk assessment. It could be argued that a state licensing agency was not upholding its gatekeeper function by allowing programs to operate with such regulatory non-compliance.

Why is the regulatory compliance data distribution important from a statistical point of view. Generally when we are dealing with social science data, the data are normally distributed or pretty close to being normally distributed. It is a trade mark of a well designed assessment tool for example. So when data are compared to other normally distributed data, there is a good chance that some form of a linear relationship will be ascertained, albeit, not reaching statistical significance in many cases but linear regardless.

When a very skewed data distribution is one of the variables as in the case with regulatory compliance data and it is compared with a normally distributed data set such as a program quality tool, ERS or CLASS. Well, the result is generally a non-linear relationship with a marked ceiling effect or plateau effect. In other words, the data distribution is more curvilinear than linear. From a practical standpoint this creates selection

problems in the inability to identify the best programs that have full regulatory compliance. This can create a public policy nightmare in that those programs which are in substantial but not full regulatory compliance are as good or in some cases of higher quality than those programs in full regulatory compliance. The interesting question is does the combination of normally distributed data distributions with variables that have skewed data distributions always produce this nonlinear result?!

And lastly, will having two variables that are skewed data distributions produce a more random result than if one of the two above conditions are present?

The Ten Principles of Regulatory Compliance Measurement

The first principle deals with the lack of Variance in data distributions. Data are found to be tightly grouped at high compliance levels (upper 90% level). This will lead to another principle addressed later in this paper dealing with skewness of the data distribution. In fact, the majority of scores are at a full regulatory compliance level, in other words, 100% in compliance with all rules and regulations. This led to variance statistics showing little movement and the majority of programs being in very close proximity. This makes for difficult statistical analyses when there is little variance in the data set.

The second principle is finding a ceiling or plateau effect in data distributions. It was like there was a diminishing returns effect as one moves from substantial regulatory compliance (upper 90%+) to full regulatory compliance (100%) with all rules and regulations. This was especially true when one compares the regulatory compliance levels with program quality scores on those same programs which is addressed more in the next principle.

The third principle is the difficulty distinguishing levels of quality between full and substantial compliance. This principle builds off of the previous principle dealing with a ceiling or plateau effect. Because so much of the data, as much as 70-80% of programs, are grouped so tightly at the substantial and high levels of regulatory compliance when one begins to go beyond regulatory compliance and begin to look at quality there is a great deal of difficulty distinguishing levels of quality. In other words, the full regulatory compliant level programs are not necessarily the highest quality programs.

The fourth principle is the fact that rules and regulations are measured at a nominal measurement level: the rules and regulations are either In-Compliance or Out-of-Compliance.

The rule or regulation is measured at a "Yes" or "No" level or a "1" or "0" level. There are no in-between measures, no ordinal measurement going on. Either you got it, or you don't. It is black or white, no shades of gray. It is just the nature of measurement when it comes to rules and regulations which are very different in other measurement systems. The data are very discrete and not continuous. They are frequency counts and not a ruler type of measurement. One will not find an interval level of measurement in any regulatory science data distribution.

A fifth principle is attempting to move to an ordinal measurement level when quality is included. This principle builds off of the previous principle in which in some cases it has been suggested to add a quality component to particular rules or regulations. This is an interesting development and moves the philosophy from one of "Do no harm" to one of "Do things well". It will be interesting to see how much this concept moves forward and changes a basic tenet in the regulatory science field which is more based upon health & safety, gatekeeper, hard data, risk aversion, and deficit based.

The sixth principle of regulatory compliance measurement is the ability to dichotomize the data can be warranted because of the data distribution. Data dichotomization is generally not recommended because it accentuates differences in a data set. However, given the nature of regulatory compliance measurement being at a nominal level, fitting into a bucket format, the lack of variance, and the skewness of the data distribution all lead to the ability to dichotomization of the data set.

The seventh principle has to do with the problem with false negatives and positives, especially false negatives. Because of the data being measured in a nominal In-Compliance vs Out-of-Compliance dichotomy it can lead to false negatives in which In-Compliance decisions are made that in reality are not In-Compliance. False positives are a problem as well but not as much of a problem as false negatives. In false positives, Out-

of-Compliance may be determined when in reality the rule or regulation is actually In-Compliance. This is not a good scenario for the provider of services, but it potentially doesn't harm the client as much as when a false negative occurs.

The eighth principle is the lack of reliability and validity testing. This principle builds from the previous principle in that there are very few examples of scientific testing of instrumentation and the administration of protocols to make certain that everything is running as it should. Because of this, it leads to the above problem of false positives and negatives. All jurisdictions need to build in regular reliability and validity testing to ascertain that the final decision making is within the ranges that are acceptable.

The ninth principle is the ease in distinguishing levels of quality between low and substantial compliance. The one result that has been consistent over the years is the ability to see differences in programs that score low on regulatory compliance versus those that are at a substantial or high compliant level. From a licensing or regulatory administration point of view this is a real plus in being able to be an effective gatekeeper and keeping non-optimal programs out of service. But as indicated in the third principle this advantage is short-lived as one moves up the regulatory compliance scale to substantial and finally to full regulatory compliance. When one gets to these levels it becomes increasingly difficult to distinguish differences in quality in those programs that are in substantial regulatory compliance versus those that are in full regulatory compliance. It appears that the regulatory compliance theory of diminishing returns is rearing its plateau/ceiling effect. The policy implications are immense since the assumption is that there is a linear relationship between program quality and regulatory compliance. How do we more effectively deal with this non-linear relationship in formulating public policy regarding licensing decision making?

And the final tenth principle is that regulatory compliance data are always skewed data. The majority of programs are in substantial or full regulatory compliance. And in many cases, this can be rather severe. There generally is a long tail which contains some low regulatory compliant programs, but these are usually few in number. The data distribution just does not approach a normally distributed curve as we see in many other examples of social science data distributions.

It is important as the regulatory science field moves forward that we remain cognizant of the limitations of regulatory compliance measurement. There are some severe limitations that need to be addressed (e.g., skewed data, lack of variance in data, ceiling effect, nominal metrics) and building in mitigation strategies (e.g., data dichotomization) or it will continue to lead to problems in our analyses (e.g., false positives and negatives).

References (https://rikinstitute.com/publications/)

(Please go to the above website if you are interested in downloading any of the publications listed here)

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

Panlilio, Famularo, Dore, Grable, Verdiglione, Yang, Lehman, Hamm, Fiene, Bard, Levi (2020). Using mixed methods to evaluate and improve the iLookOut for Child Abuse knowledge scale based on structural and response process validity evidence, Early Childhood Research Quarterly.

Levi, Belser, Kapp, Verdiglione, Mincemoyer, Dore, Keat, Fiene (2020). An Interactive Online Learning Program on Child Abuse and Its Reporting, Journal of Early Childhood Teacher Education.

Kapp, Dore, Fiene, Grable, Panlilio, Hamm, Yang, Lehman, Mincemoyer, Verdiglione, Levi (2020). Cognitive Mapping Utilized With An Online Training Program: Developing A Conceptual Framework For Learning and Performance Outcomes in An Online Learning Program, Journal of Distance Education and e-Learning, Volume 8, Number 2.

Yang, Panlilio, Lehman, Hamm, Fiene, Verdiglione, Barr, Grable, Levi (2020). iLookOut phase II: Generalized findings from a randomized controlled trial to a real world study of an online education program on child abuse reporting, Plos One

Public Library of Science, PLoS ONE 15 (1): e0227398. (doi.org/10.1371/journal.pone.0227398).

Fiene, Richard (2019), "ECPQIM National Data Base", Mendeley Data, v2, (dx.doi.org/10.17632/kzk6xssx4d.1).

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

Panlilio, Yang, Verdiglione, Fiene, Lehman, Hamm, McNeil, & Levi (2018). Effectiveness of an online education program to improve knowledge and attitudes about mandated reporting for early childhood care and education providers, National Research Conference on Early Childhood, Washington DC.

Johnston, DelConte, Aronson, & Fiene (2017). Caring for our children standards into child care practice: child care health consultation improves infant and toddler care, manuscript prepared for submission to Journal of Pediatric Health Care, Bryn Mawr, PA: Pa Chapter of the American Academy of Pediatrics. (DOI: 10.13140/RG.2.2.10119.80808)

Fiene (2016). Early Childhood Program Quality Improvement and Indicator Model and the Differential Monitoring Logic Model and Algorithm Book of Readings. Middletown, PA.: Research Institute for Key Indicators.

Fiene & Kroh (2016). Licensing measurement, regulatory compliance and program monitoring systems, in Licensing Curriculum, Lexington, KY: National Association for Regulatory Administration.

Fiene, (2015). Early childhood program quality indicators model (Version 4): Differential monitoring, risk assessment

and key indicators, Middletown, PA.: Research Institute for Key Indicators.

Lahti, Elicker, Zellman, & Fiene (2014). Approaches to validating child care quality rating and improvement systems (QRIS): Results from two states with similar QRIS type designs, Early Childhood Research Quarterly, available online 9 June 2014, doi:10.1016/j.ecresq.2014.04.005.

Fiene (2013). A Comparison of International Child Care and US Child Care Using the Child Care Aware – NACCRRA (National Association of Child Care Resource and Referral Agencies) Child Care Benchmarks, International Journal of Child Care and Education Policy, 7(1), 1-15.

Zellman & Fiene (2012). Validation of quality rating and improvement systems for early care and education and school-age care, Washington, D.C.: OPRE and Child Trends.

Fiene & Carl (2011). Child Care Quality Indicators Scale, in T Halle (Ed.), Quality Rating and Improvement Systems Tool Kit, Washington, D.C.: Child Trends.

Manlove, Benson, Strickland, & Fiene (2011). A comparison of regulated child care in rural and urban Pennsylvania, Pennsylvania: Center for Rural Pennsylvania.

Johnson, Fiene, McKinnon, & Babu (2010). A study of early childhood education pre-service teacher education at major universities in 38 pre-k states. New York, New York: Foundation for Child Development.

Fiene (2009). Mentoring programs help caregivers “weigh” in on children’s eating behaviors, Penn State Policy Briefing, University Park, Pennsylvania: Center for Public Policy Research in Environment, Energy and Community.

Fiene (2008). Using Child Care Programs As A Portal to Changing the Eating Behaviors of Young Children, in L. Birch and W. Dietz (Eds.), Eating Behaviors of the Young Child: Prenatal and Postnatal Influences on Healthy Eating, Elk Grove Village, Illinois: American Academy of Pediatrics.

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 and Isler (2007). Home based and family child care: Characteristics and quality issues, In C. J. Groark, K. E., Mehaffie, R. B. McCall & M. T. Greenberg (Eds.), Evidence-Based Programs, Practices and Policies for Early Childhood Care and Education. Thousand Oaks, CA: Corwin Press.

Weinraub, Shlay, Kochanoff & the Universities Children's Policy Collaborative (Fiene, Greenberg, McCall, Groark, Mehaffie, Nelkin & Ethridge-Smith) (2006). Findings from the 2002 Pennsylvania Family Survey and their relevance for the future, Commonwealth, Pennsylvania House of Representatives (LORL).

Fiene (2006). Early Childhood Initiatives in Other States, in J.M. Perzel, Education in Pennsylvania: Early Childhood Education: Universal Pre-K and Other Alternatives, Volume 4, Spring, 2006, Harrisburg: Pennsylvania House of Representatives.

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

Fiene, Greenberg, Bergsten, Carl, Fegley, & Gibbons (2002). The Pennsylvania early childhood quality settings study,

Harrisburg, Pennsylvania: Governor's Task Force on Early Care and Education.

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

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.

Johnson, Fiene, Keat, Darling, Pratt, Iutcovich (2001). Mastering course content and learner satisfaction: A comparison of regular classroom instruction with three variations of internet delivery. Journal of Early Childhood Teacher Education, 22(4), 267-274.

Iutcovich, Fiene, Johnson, Koppel, & Langan (2001). Professional development and the quality of child care: An assessment of Pennsylvania's child care training system. In Early education and care, and reconceptualizing play, Elsevier Science Ltd., Volume 11, 115-168.

Fiene, & Kroh (2000). Measurement tools and systems, in Licensing Curriculum, National Association for Regulatory Administration, Minneapolis, Minnesota.

Fiene, Iutcovich, Johnson, & Koppel (1998). Child day care quality linked to opportunities for professional development: An applied community psychology example. Community Psychologist, 31(1), 10-11.

Fiene, (1997). Quality child care in Pennsylvania linked to opportunities for professional development. National

Association of Regulatory Administration, Autumn 1997, page 13.

Fiene (1997). Potential solution to the child day care trilemma related to quality, accessibility and affordability. Child Care Information Exchange, September, 57-60.

Fiene (1997). Human services licensing information system. National Association for Regulatory Administration: Research Column, Spring, 9-10.

Iutcovich, Fiene, Johnson, Koppel, & Langan (1997). Investing in our children's future, Erie, Pennsylvania: Keystone University Research Corporation.

Fiene (1996). Unannounced versus announced licensing visits in monitoring child day care programs. National Association of Regulatory Administration, Spring, 5-6.

Fiene (1996). The feasibility of using sampling of agency facilities. National Association for Regulatory Administration Licensing Newsletter, fall, 12-14.

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

Fiene (1995). Utilizing a statewide training system to improve child day care quality: The other system in a program quality improvement model. Child Welfare, Volume LXXIV, #6, November-December, 1189-1201.

Fiene (1995). Improving human services through licensing reform and training: A unique partnership between academia and state government, Dimensions, Fall, 1-3.

Kuhns & Fiene (1995). Promoting health and safety in child care programs, Child Care Bulletin, January-February (1), 3.

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

Fiene (1994). The case for national early care and education standards: Key indicator/predictor state child care regulations, National Association of Regulatory Administration, summer 1994, 6-8.

Fiene (1991). New early childhood research, evaluation and training program has impact on Pennsylvania for the 1990's, Dimensions, Fall, 4.

Fiene (1988). Human services instrument based program monitoring and indicator systems, in Information Technology and the Human Services, B. Glastonburg, W. LaMendola, & S. Toole, editors, Chichester, England: John Wiley and Sons.

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

Fiene (1987). Using licensing data in human service programs, in Licensing, H. Hornby, editor, Portland, Maine: University of Southern Maine.

Fiene (1987). The indicator system, in Evaluation and outcome monitoring, H. Hornby, editor, Portland, Maine: University of Southern Maine.

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

Fiene (1987). Indicator checklist system, in Maximizing the Use of Existing Data Systems, Portland, Maine: University of Southern Maine.

Fiene (1986). State child care regulatory, monitoring and evaluation systems as a means for ensuring quality child development programs, in Licensing of Children's Services Programs, Richmond, Virginia: Virginia Commonwealth University School of Social Work. (ERIC/ECE ED322997)

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.

Kontos & Fiene (1986). Predictors of quality and children's development in day care, in Licensing of Children's Services Programs, Richmond, Virginia: Virginia Commonwealth University School of Social Work.

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

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

Fiene & Nixon (1983). Indicator checklist system for day care monitoring, Washington, D.C.: National Children's Services Monitoring Consortium.

Fiene & Nixon (1981). An instrument-based program monitoring information system: A new tool for day care monitoring, Washington, D.C.: National Children's Services Monitoring Consortium.

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

Fiene (1979). Potpourri of child development, In the Best Interests of Children, Spring, 14-15.

Fiene (1978). Theoretical model for computing adult child ratios, Association of Regulatory Administration, summer, 12-13.

Fiene (1975). Current trends in child day care research, In the Best Interests of Children, Spring, 15-18.

Fiene, Cardiff, & Littles (1975). Ecological monitoring information system, In the Best Interests of Children, July-September, 1975.

Fiene (1974). The two year old: Characteristics and management of his play, Dimensions: Journal of the Southern Association of Children Under Six, January, 1974, 2(2), 46-48.

About the Author:

After a long career in governmental service and academia, mostly in Pennsylvania; and consulting, nationally and internationally, Dr Rick Fiene continues to write and research about regulatory science topics (such as measurement, instrument development, math & statistical modeling, differential monitoring, risk assessment, key performance indicators) as they related to early care and education, the human services, and has been delving into other social sciences as well.

https://www.prevention.psu.edu/people/rick-fiene

Contact Dr Fiene at his Research Institute:

rfiene@rikinstitute.com

or through the National Association for Regulatory Administration:

rfiene@naralicensing.org

or through the Penn State Prevention Research Center:

fiene@psu.edu

RESEARCH INSTITUTE FOR KEY INDICATORS

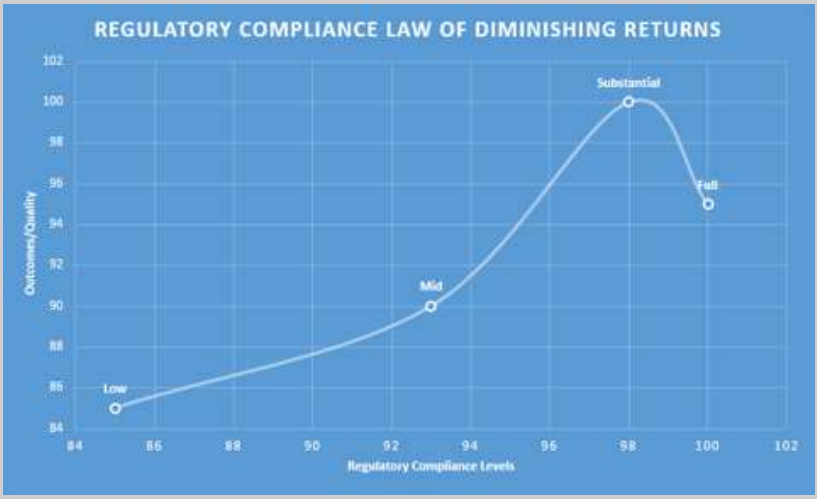
NATIONAL ASSOCIATION FOR REGULATORY ADMINISTRATION

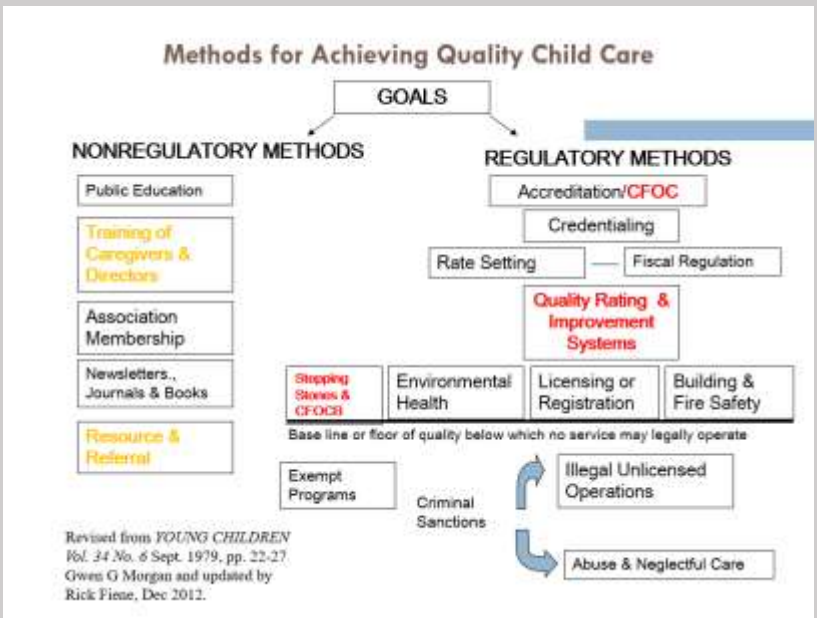
THE PENNSYLVANIA STATE UNIVERSITY

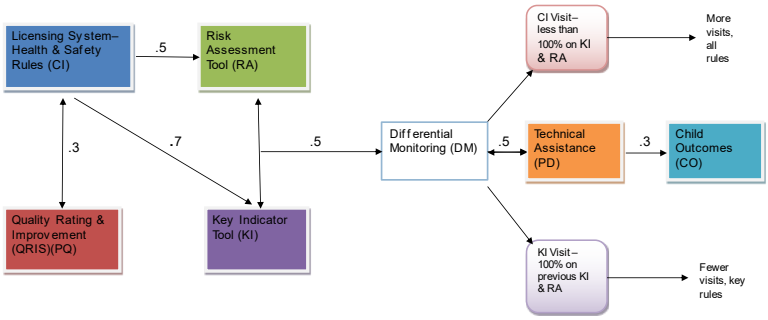
The below graphics, figures and displays help to support and depict various portions of text in Chapters 1-7.

They depict the regulatory compliance theory of diminishing returns, the overall regulatory compliance and program quality model as described by Gwen Morgan, the ECPQIM: Early Childhood Program Quality Improvement and Indicator Model that depicts the relationship of regulatory compliance and program quality monitoring systems; a brief logic model of when risk assessment and key indicator methods can and cannot be used; the RAM: Risk Assessment Matrix decision matrix; the relationship of comprehensive reviews and abbreviated reviews, such as, risk assessment and key indicator reviews; data distributions for regulatory compliance, and program quality as depicted with QRIS and ERSs; ECPQIM theory; Key indicator and non-compliance relationship depicting the relationship between effectiveness and efficiency; the key indicator and risk assessment methodologies within a single matrix format; the use of Caring for Our Children in depicting the relationship between compliance and quality; the Regulatory Compliance Scale; the absolute and differential regulatory paradigms key elements; International study of child care comparing rules and regulations; Key indicator statistical formula; examples of two data distributions from Head Start (skewed) and ECERS (normally distributed); ECPQIM version 5 which demonstrates the use of integrated monitoring.

Figures, Graphs, Charts, Displays



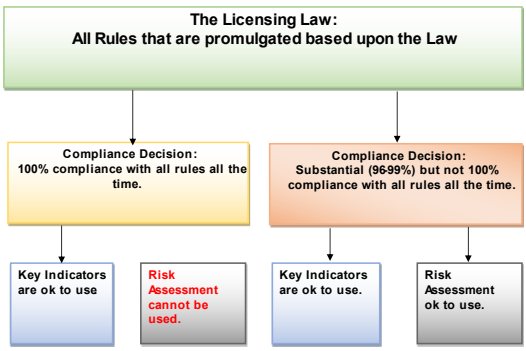




$$\sum CI \times \sum PQ \Rightarrow \sum RA + \sum KI \Rightarrow \sum DM + \sum PD \Rightarrow CO$$

29

When Key Indicators and Risk Assessments Can Be Used



40

Risk Assessment Matrix (RAM)

Risk Assessment (RA) Matrix Revised

Risk/Severity Levels	High	Medium	Low
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.	4-7 rules out of compliance. 93 – 97 regulatory compliance.	3 or fewer rules out of compliance. 98 – 99 regulatory compliance.

***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

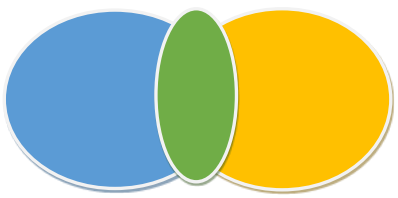
Tier 3 = ((RC = 93 – 97) + (Medium Risk)); ((98 – 99) + (Medium Risk)) = Tier 3

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+

Using RAM to make licensing decisions

Relationship of Comprehensive Reviews (CR) to Key Indicator (KI) or Risk Assessment (RA) Rule Non - Compliance

Key Indicator Rule | Both | Risk Assessment Rule



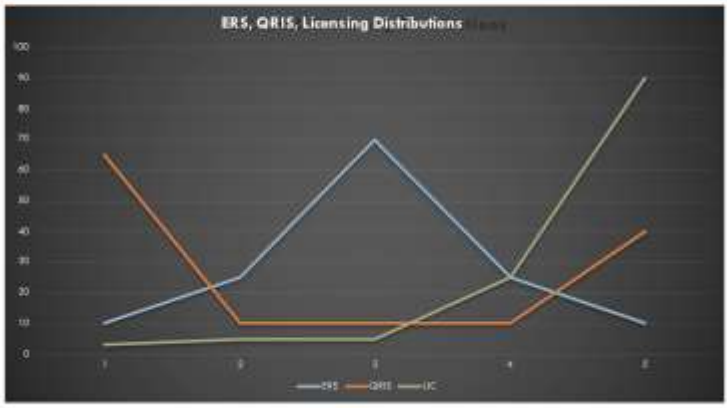
Prediction

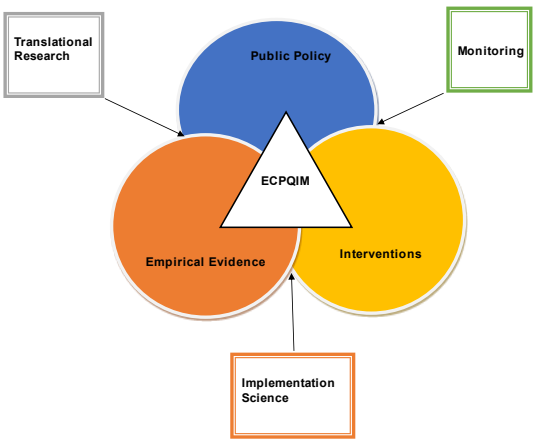
Non-Compliance	***Non-Compliance***	***Non-Compliance***
2+ Rules = CR	1 Rule = CR	Point System = CR
1 Rule = Section		1 Extreme Rule = CR
Absolute scoring 1/0		Relative scoring 1/9

Risk to Children

ERS, QRIS, Licensing Comparisons

112





137

Key Indicator/Non-Compliance Relationship

84

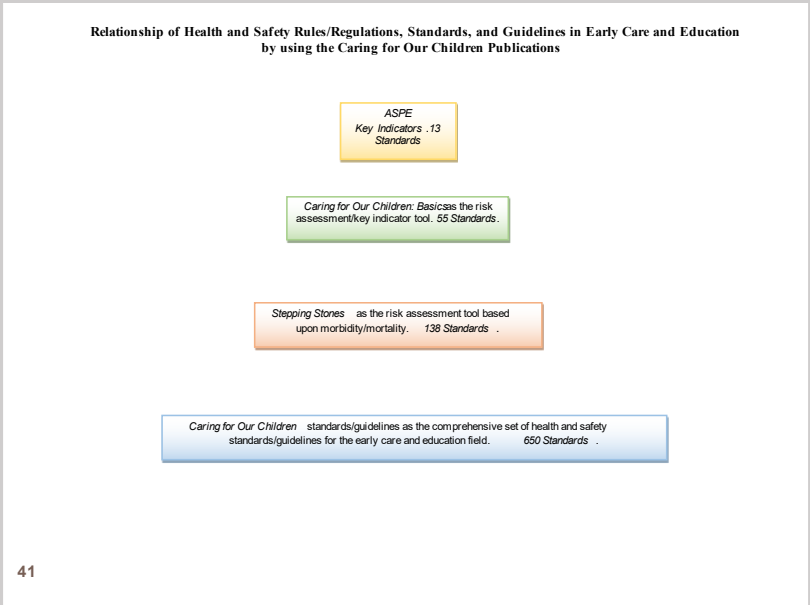
KIM (Key Indicator Matrix) and RAM (Risk Assessment Matrix) Matrices Integration Into One Platform

83

KIM	Low Group	High Group		Severity:
Compliance	1	2	3	Low
Non-Compliance	4	5	5	Medium
	7	8	9	High
Prevalence:	Low	Medium	High	RAM

This technical research note will integrate the Key Indicator Matrix (KIM) and the Risk Assessment Matrix (RAM) into one platform to clearly demonstrate their statistical modeling overlap. Key Indicators deal with the ability to predict overall compliance or performance based on existing data. Risk Assessment Indicators do not predict but determine a risk score based upon prevalence and severity measures. Their purposes are different but when integrated together the two matrices are a powerful tool in determining the health of the measured entity.

The above matrix integrates the two matrices of KIM and RAM and shows that KIM scores are generally at the lower end of risk but having sufficient prevalence when it comes to non compliance. RAM scores have a larger variance and are most concerning at the higher end of the continuum



Regulatory Compliance Scale (RCS)

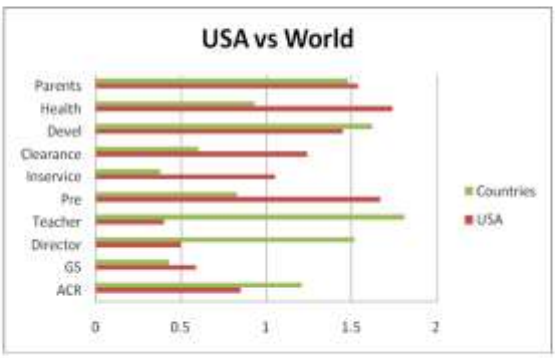
37

Regulatory Compliance Scale Levels	Definitions & Compliance Levels	Number of Rule Violations
7	Full 100% Compliance	0 Violations
5	Substantial Compliance	1-3 Violations
3	Mediocre Compliance	4-9 Violations
1	Low/Non-Optimal Compliance	10+ Violations

Regulatory Paradigms

Absolute (Class, 1957)	Relative/Differential (Fiene, 1985)
All rules are created equal.	All rules are not created equal.
100% Compliance = Full License.	Full 100% + Substantial Compliance = Full License.
PC + PQ = Linear.	PC + PQ = Not Linear.
All rules are reviewed all the time.	Selected key rules are reviewed all the time.

International Study of Child Care Rules (Fiene, 2013a)



Key Indicator Statistical Methodology

87

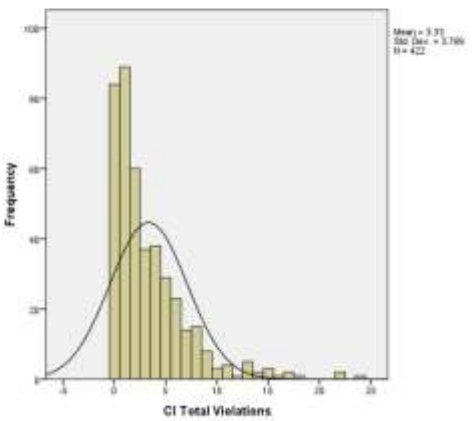
$$\phi = (A)(D)-(B)(C) \div \sqrt{(W)(X)(Y)(Z)}$$

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.

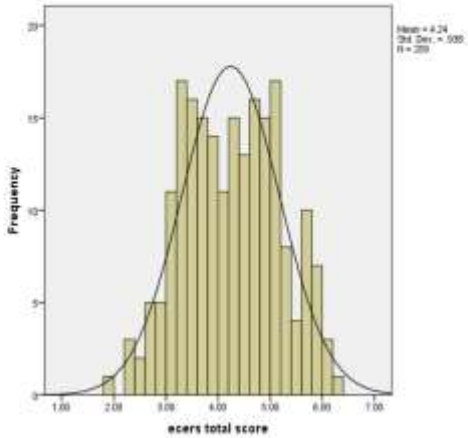
Head Start Performance Standards

111



ECERS Total Scores

108



ECPQIM5: Early Childhood Program Quality Improvement/Indicator Model Version 5